LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-413970

Cable Damage Detection System and Algorithms Using Time Domain Reflectometry

G. A. Clark, C. L. Robbins, K. A. Wade, P. R. Souza

June 16, 2009

Disclaimer

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Cable Damage Detection System and Algorithms Using Time Domain Reflectometry

Grace A. Clark, Ph.D., IEEE Fellow
Lawrence Livermore National Laboratory
7000 East Ave., L-130, Livermore, CA 94550
(925) 423-9759 (Office), (925) 422-2495 (FAX),
clark9@llnl.gov

Christopher L. Robbins
Lawrence Livermore National Laboratory
7000 East Ave., L-153, Livermore, CA 94550
(925) 422-2274 (Office),
robbins12@llnl.gov

Katherine A. Wade
Lawrence Livermore National Laboratory
7000 East Ave., L-153, Livermore, CA 94550
(925) 422-1567 (Office),
wade15@llnl.gov

Paul R. Souza
Lawrence Livermore National Laboratory
7000 East Ave., L-093, Livermore, CA 94550
(925) 423-4862 (Office)

June 12, 2009

Acknowledgments

The TDR project team is multidisciplinary. Christopher Robbins served as project manager, conducted the test program to acquire data and instrument the process and contributed greatly to the repeatability software and analysis. Dr. Grace Clark served as a specialist in signal processing/estimation/detection. She wrote algorithm plans, test plans and signal processing algorithm software. Katherine Wade wrote overall software tools and graphical user interfaces for the signal processing software. In particular, she built receiver operating characteristic (ROC) curves and confidence intervals for the cable damage tests. Paul Souza and Trung Le expertly conducted laboratory experiments and acquired the data. Rex Morey was the initial principal investigator for the time domain reflectometry effort. Eric Breitfeller provided signal processing and software engineering. Dr. Sean K. Lehman provided an initial analysis of repeatability statistics. Jessie A. Jackson built the initial repeatability test software and GUI. Dr. Graham H. Thomas has served as program leader for the ADAPT Life Extension Program. Karl Krause served as program leader for the ADAPT program and Derek Wapman served as program leader for the W80 program. Tim Rau served as W80 program leader during 2006.

Abstract

This report describes the hardware system and the set of algorithms we have developed for detecting damage in cables for the Advanced Development and Process Technologies (ADAPT) Program. This program is part of the W80 Life Extension Program (LEP). The system could be generalized for application to other systems in the future.

Critical cables can undergo various types of damage (e.g. short circuits, open circuits, punctures, compression) that manifest as changes in the dielectric/impedance properties of the cables. For our specific problem, only one end of the cable is accessible, and no exemplars of actual damage are available. This work addresses the detection of dielectric/impedance anomalies in transient time domain reflectometry (TDR) measurements on the cables. The approach is to interrogate the cable using time domain reflectometry (TDR) techniques, in which a known pulse is inserted into the cable, and reflections from the cable are measured. The key operating principle is that any important cable damage will manifest itself as an electrical impedance discontinuity that can be measured in the TDR response signal. Machine learning classification algorithms are effectively eliminated from consideration, because only a small number of cables is available for testing; so a sufficient sample size is not attainable. Nonetheless, a key requirement is to achieve very high probability of detection and very low probability of false alarm.

The approach is to compare TDR signals from possibly damaged cables to signals or an empirical model derived from reference cables that are known to be undamaged. This requires that the TDR signals are reasonably repeatable from test to test on the same cable, and from cable to cable. Empirical studies show that the repeatability issue is the "long pole in the tent" for damage detection, because it is has been difficult to achieve reasonable repeatability. This one factor dominated the project. The two-step model-based approach is summarized as follows:

Step 1, Cable Modeling: Given input-output TDR signals $s(n)$ and $x_U(n)$ for a cable known to be free of damage, system identification algorithms are used to compute a dynamic prediction-error cable model that has output $\widehat{x}_U(n)$. The model is declared valid when the innovations $e_U(n) = x_U(n) - \widehat{x}_U(n)$ satisfy a statistical zero-mean whiteness test. This validated model output is then used as a known reference to which other cables can be compared.

Step 2, Cable Testing: The TDR output signal $x_D(n)$ from a cable under test is compared with the model output $\widehat{x}_U(n)$ by computing the innovations $e_D(n) = x_D(n) - \widehat{x}_U(n)$. The innovations are tested using a short-term whiteness test statistic, which employs a statistical confidence interval. If the cable passes the test, this implies that the model is valid and the cable is declared undamaged. If the cable fails the test, this indicates a model mismatch, which means that the cable's dielectric properties have changed; and this implies that the cable is damaged. The test threshold is adjusted to maximize probability of detection and minimize probability of false alarm according to an empirically determined receiver operating characteristic (ROC) curve. An associated confidence interval on the probability of correct classification is also provided.

The effectiveness of the algorithm is demonstrated using measured TDR signals for undamaged and damaged cables. Experimental and algorithmic methods for coping with repeatability issues are presented.

The model-based damage detection algorithms have been shown to perform well for some representative examples of real TDR signals acquired using the two-dimensional (2D) mockup fixture. If the damage causes a short circuit, then damage detection performance is generally good to excellent. Examples include the cases demonstrated in this report for cuts and pinholes. If the damage does not cause a short circuit, then damage detection performance is generally poor to fair. Examples include the cases demonstrated in this report for crushes and partial cuts.

Contents

List of Figures

Chapter 1

Introduction and Problem Definition

This report describes the hardware system, the set of algorithms we have developed for detecting damage in cables, and processing results for the Advanced Development and Process Technologies (ADAPT) Program. This program is part of the W80 Life Extension Program (LEP). The system could be generalized for application to other systems in the future.

Critical cables can undergo various types of damage (e.g. short circuits, open circuits, punctures, compression) that manifest as changes in the dielectric/impedance properties of the cables. For our specific problem, only one end of the cable is accessible, and no exemplars of actual damage are available. This work addresses the detection of dielectric/impedance anomalies in transient time domain reflectometry (TDR) measurements on the cables. The approach is to interrogate the cable using time domain reflectometry (TDR) techniques, in which a known pulse is inserted into the cable, and reflections from the cable are measured. The key operating principle is that any important cable damage will manifest itself as an electrical impedance discontinuity that can be measured in the TDR response signal. Machine learning classification algorithms are effectively eliminated from consideration, because only a small number of cables is available for testing; so a sufficient sample size is not attainable. Nonetheless, a key requirement is to achieve very high probability of detection and very low probability of false alarm.

The approach is to compare TDR signals from possibly damaged cables to signals or an empirical model derived from reference cables that are known to be undamaged. This requires that the TDR signals are reasonably repeatable from test to test on the same cable, and from cable to cable. Empirical studies show that the repeatability issue is the "long pole in the tent" for damage detection, because it is has been difficult to achieve reasonable repeatability. This one factor dominated the project. The two-step model-based approach is summarized as follows:

Step 1, Cable Modeling: Given input-output TDR signals $s(n)$ and $x_U(n)$ for a cable known to be free of damage, system identification algorithms are used to compute a dynamic prediction-error cable model that has output $\widehat{x}_U(n)$. The model is declared valid when the innovations $e_U(n) = x_U(n) - \widehat{x}_U(n)$ satisfy a statistical zero-mean whiteness test. This validated model output is then used as a known reference to which other cables can be compared.

Step 2, Cable Testing: The TDR output signal $x_D(n)$ from a cable under test is compared with the model output $\widehat{x}_U(n)$ by computing the innovations $e_D(n) = x_D(n) - \widehat{x}_U(n)$. The innovations are tested using a short-term whiteness test statistic, which employs a statistical confidence interval. If the cable passes the test, this implies that the model is valid and the cable is declared undamaged. If the cable fails the test, this indicates a model mismatch, which means that the cable's dielectric properties have changed; and this implies that the cable is damaged. The test threshold is adjusted to maximize probability of detection and minimize probability of false alarm according to an empirically determined receiver operating characteristic (ROC) curve. An associated confidence interval on the probability of correct classification is also provided.

The effectiveness of the algorithm is demonstrated using measured TDR signals for undamaged and damaged

cables. Experimental and algorithmic methods for coping with repeatability issues are presented.

1.1 The Electronic Cable Assembly Tester System (ECATS)

Our overall goal is in situ reliability testing of cables installed in a hardware system. We are interested in detecting any possible damage that could prevent the cable from functioning properly. For our problem, only one end of the cable is accessible. The approach is to interrogate the cable using time domain reflectometry (TDR) techniques, in which a known pulse is inserted into the cable, and reflections from the cable are measured. The key operating principle is that any important damage will manifest itself as an electrical impedance discontinuity that can be measured in the TDR response signal.

Cables suffer insults during fabrication and installation and from environmental stresses and aging. Simple testing tools can determine cable continuity, but they are not enough. There are many situations where cable integrity is compromised and will fail operationally or gradually over time, yet continuity is not broken. In these situations, existing testing tools, such as the ubiquitous volt-ohm meter, are useless.

The Electromagnetic Cable Assembly Tester System (ECATS) was developed to measure cable integrity in high-value assets where reliability is paramount and failure could be damaging or even catastrophic, and where no other diagnostic tool is available. ECATS measures the integrity of an embedded cable assembly with unprecedented accuracy and in an operational environment without damaging or compromising the system. The Time Domain Reflectometry (TDR) measurement system is summarized in LLNL reports [10, 11].

1.2 The Cables and TDR

The ECATS measurement system is summarized in Figure (1.1). The pulse insertion unit (PIU) uses capacitive coupling and is designed to match the impedance of the cable under test to the impedance of the TDR measurement system. See the chapter on ECATS for more details. The cable of interest is a two-conductor flat cable with Kapton insulation. Figure (1.2) depicts a sample cable and example TDR signals for the two cases in which the cable is undamaged and damaged. The figure depicts a comparison of two cable TDR signals; one for an undamaged cable and one for a cable with simulated damage. The damage was simulated by putting a strip of foil around an undamaged cable to produce a dielectric/impedance discontinuity, as described in Figure (1.2). The figure demonstrates that a reflection caused by the dielectric discontinuity is seen in the TDR signal for a damaged cable, but not for an undamaged cable.

1.3 The Model-Based Approach

The technical challenges are many and difficult, as summarized in Figure (1.4). Because of the small statistical sample size available (number of cables for testing and number of signals for each cable), it was decided to not use a machine learning approach. Instead we adopted a model-based approach in which we compare TDR signals from possibly damaged cables to signals or an empirical model derived from reference cables that are known to be undamaged. This requires that the TDR signals are reasonably repeatable from test to test on the same cable, and from cable to cable. With enough resources, it might be possible to detect damage, classify the damage type, locate the damage decide whether or not the cable is reliable in some sense. See Figure (1.6). However, due to resource constraints, we focused on the detection problem, in which the goal is to decide whether a dielectric anomaly is present in the cable under test. Such a decision would imply that the cable is damaged and the cable should be inspected further.

The model-based approach has the following advantages: (1) It can be implemented with a minimum number

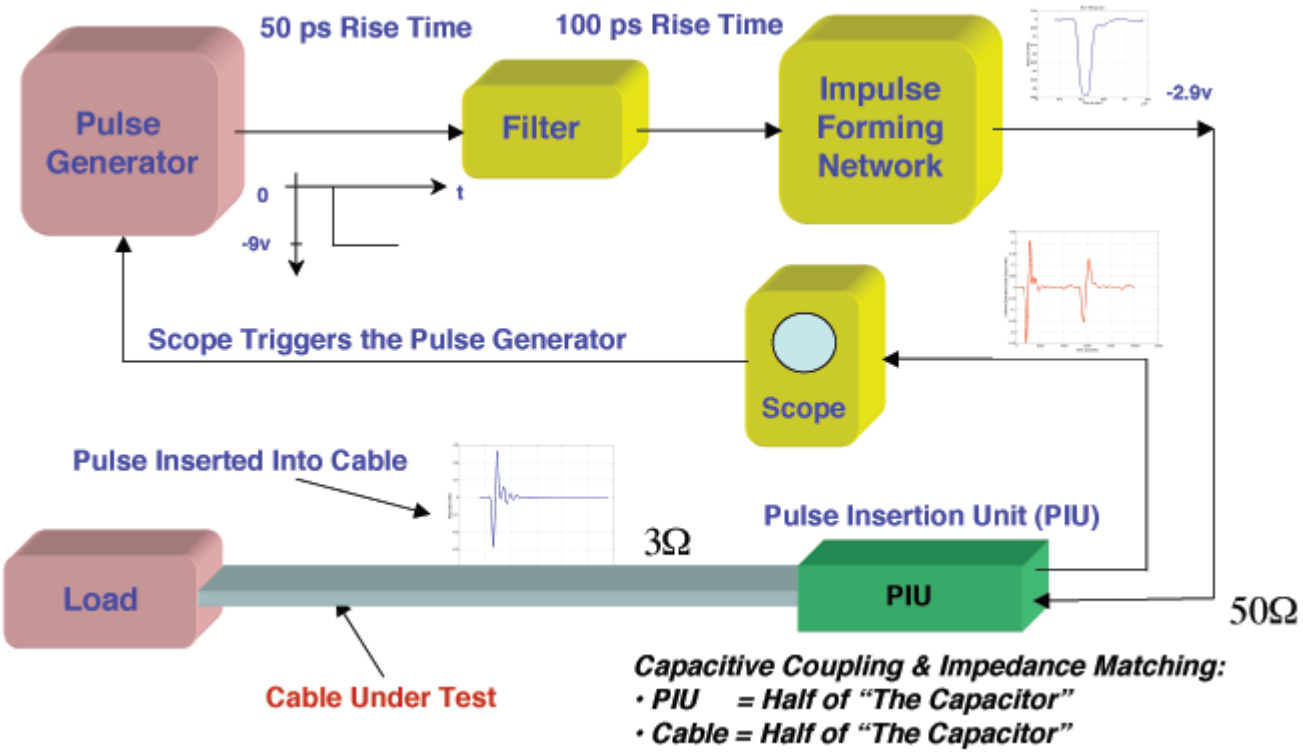


Figure 1.1: The key hardware component of the TDR system is the Pulse Insertion Unit (PIU).

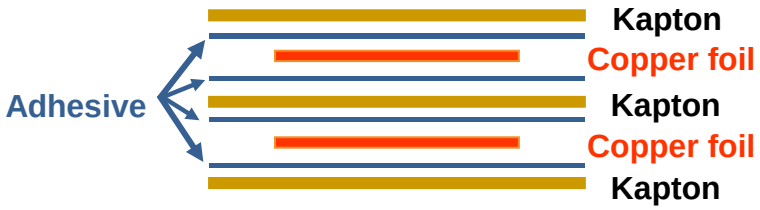


Figure 1.2: We are testing two-conductor flat cables with Kapton insulation for dielectric anomalies. This figure depicts a cross-section of a cable, showing its layers.

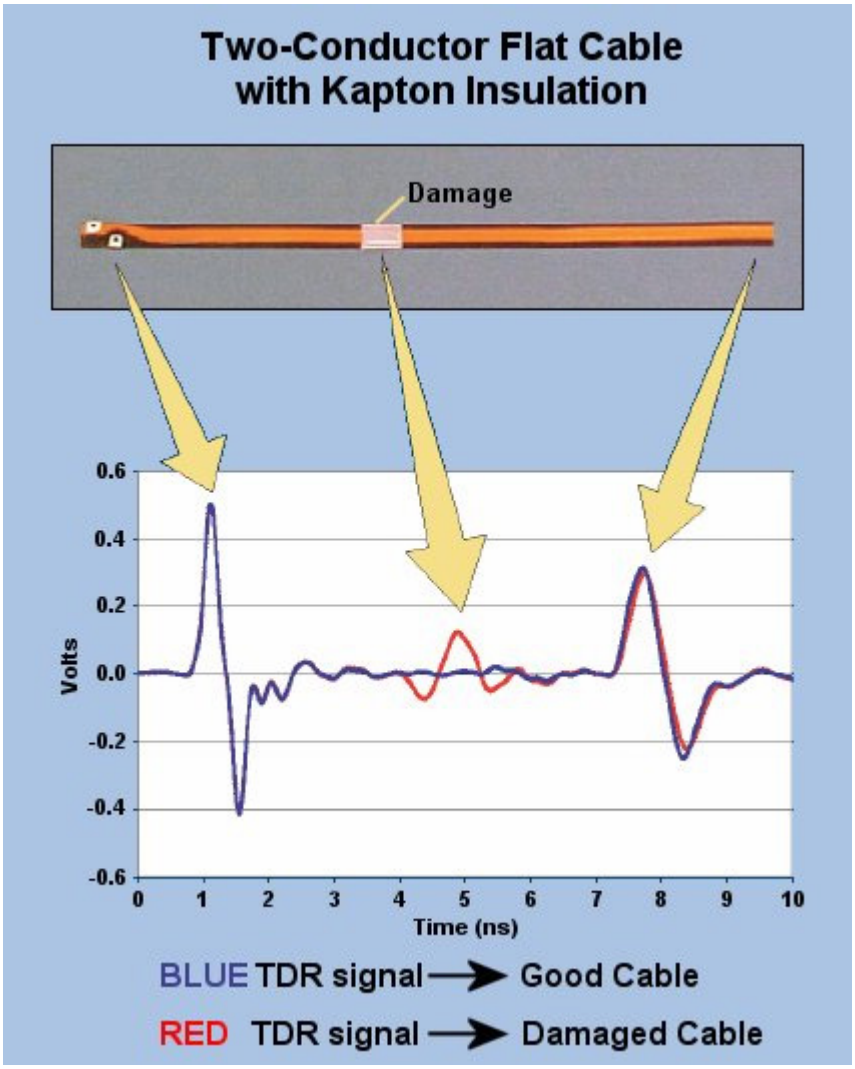


Figure 1.3: This figure expands upon the information in the last figure (Figure (1.2)). Here, one can clearly see the relationships among the positions on the cable and the reflection wavelets in the TDR signal. This figure is used for instructive purposes to demonstrate the principles of TDR. At this time, test fixtures have been fabricated and tested. The repeatability of the measurements in the fixtures has been demonstrated to be reasonable.

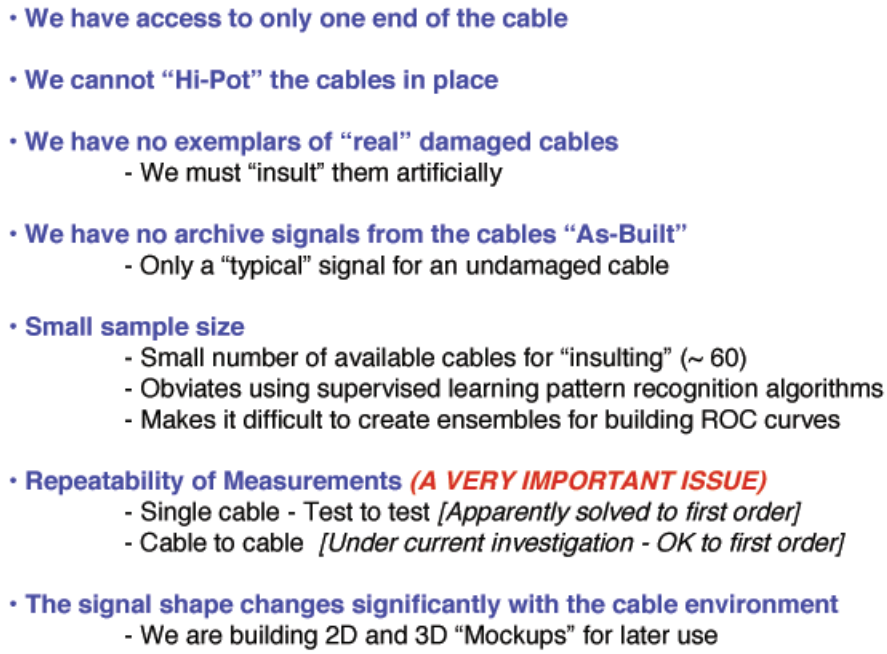


Figure 1.4: The technical challenges are difficult.

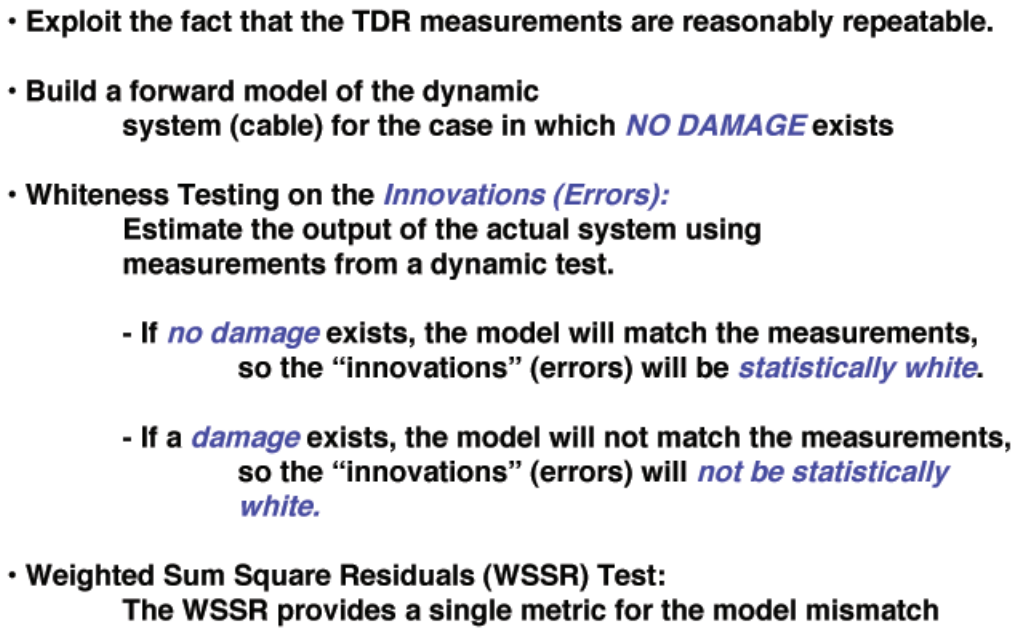


Figure 1.5: The model-based damage detection approach: Detect a model mismatch if damage is present.

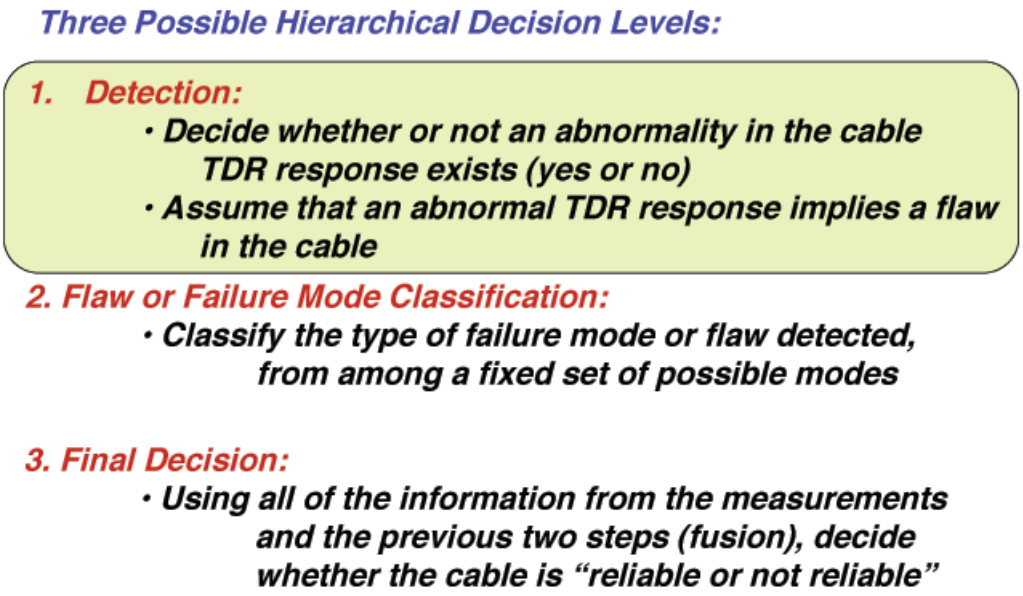


Figure 1.6: Our focus is on a binary detection decision (Yes/No), not failure mode classification or reliability.

of reference signals. (2) If the reference signal from a reference undamaged cable changes due to the physical environment in which it is placed, etc., a new model can simply be identified, without the need for a large number of new tests and classifier training steps. This is important, because the cable environment can have a profound effect on the TDR signals measured. (3) The approach can often provide information that can help locate the damage.

Chapter 2

The Measurement System

2.1 Summary of the ECATS Measurement System

A time domain reflectometry (TDR) system was developed to acquire the signals needed for damage detection. The system has been dubbed "ECATS, " for the Electromagenetic Cable Assembly Tester System [10, 11].

ECATS safely measures the integrity of transmission line systems such as actuators, electrical sensors in aircraft fuel tanks, or detonators embedded in very sensitive environments. ECATS was originally developed to meet the rigorous safety demands of high-value national security systems where safe-tester operation is of primary importance. However, ECATS is applicable to a wide range of military and commercial transmission line environments where the integrity of transmission lines is critical. Examples include satellites, aircraft, medical devices, industrial control systems, and computers.

A simple battery-operated volt-ohm meter will detect an open or short circuit in many transmission line configurations, but that is not sufficient. In contrast, ECATS provides unprecedented quantitative data about the condition of the transmission line along its full length. The hypothesized damage types include punctures, cuts, abrasions, crushes, etc. Embedded transmission lines may be damaged during assembly, or from demanding operational conditions, or long-term aging can make it susceptible to operational failure. Current inspection and maintenance techniques may not detect these problems until the cable fails, potentially leading to catastrophic failure. Transmission lines come in a large variety of geometric configurations and impedance characteristics. These include 50-ohm coaxial cable, flat low-impedance cable, and integrated circuit microstrip. ECATS modular design can accommodate these different types of transmission lines. ECATS comprises a transmission line coupling unit and a battery-operated tester connected to a laptop computer through a fiber optic cable. Depending on the transmission lines geometry and impedance of the device under test (DUT), individual coupling units are tailored to maximize signal transfer between the tester unit and the DUT, thus optimizing the signal-to-noise ratio and minimizing the amount of energy needed to interrogate the sensitive DUT. A multitude of safety layers insures even an improbable tester failure will not damage to the DUT.

ECATS is based on Time Domain Reflectometry (TDR). Very-low-energy pulses are AC-coupled into the cable under test and reflections from defects are measured with picosecond time-resolution. The resulting waveform is analyzed for both the presence and the position of a defect. The operator is alerted if the analysis indicates cable failure or the possibility of a future failure.

2.2 Key Specifications and Safety Considerations

There are six levels of protection between the DUT and power sources:
1. AC coupling to the DUT
   a. No DC path to the DUT
   b. Shorted transmission line hardware limits the TDR signal to less than 300 picoseconds.
2. Both the source and sampler circuits have resistive current limiters
3. No electrostatic discharge (ESD) can develop to DUT
   a. Transmitter (pulse generator) is an inductor-to-ground topology, so there is no capacitive storage
   b. Sampling diodes are self-biased (no DC bias)
4. Output current is limited
   a. The worst case current in the DUT is less than 50 mA RMS.
   b. Internal DC current is limited to 10 mA to all critical circuit blocks.
   c. Operation from an energy-limited source, four 1.2V AA batteries
5. AC power is isolated
   a. Fiber optic cable isolates the tester from the laptop computer to preclude DUT connection to house power sources (i.e., 110 volts ac) and potential lightning strikes.
6. Built-in self-test of current limiters on power up

See [8] for a detailed safety analysis.

2.3 Time Domain Reflectometry Principles

The primary measurement method of ECATS is based on time domain reflectometry (TDR), first developed in the 1950s by Tektronix and Hewlett Packard (now Agilent Technologies) as a laboratory 50-ohm cable evaluation instrument. Since then, TDR has become a well-established transmission line analysis technique. TDR is sometimes called closed-loop radar because it includes a transmitter and a receiver, and propagation is through a closed transmission line rather than free-space.

Commercially available TDR instruments use a very-wide bandwidth equivalent sampling oscilloscope (18-20 GHz or even more) with an internal step function generator. Connection to the Device Under Test (DUT) is via cables, probes, and fixtures. The fast rise-time step function is delivered to the DUT, and failure analysis is performed on the reflection from the DUT. Direct electrical contact between the TDR instrument and the DUT is required to perform the measurement. In addition, the ground contact must also be provided in order for the TDR signal to produce meaningful information about the DUT. Figure (2.1) shows an Agilent Technologies TDR system.

ECATS consists of the battery-operated tester unit (PT4194), transmission line coupling pulse insertion unit (PIU), and laptop computer for control, analysis, and data storage as shown in Figure (2.2). Figure (2.3) shows a series of photographs of the pre-production TDR tester. Note that the tester is compartmentalized and battery operated as per critical safety requirements.

Figure (2.4) shows the pulse insertion unit (PIU) connected to a low impedance flat cable and the TDR tester with the cover removed. Figure (2.5) is a block diagram of the TDR pulser, which is attached to the DUT with the PIU. The critical high-frequency microwave components are the impulse generator (transmitter or source) and the sampler (receiver or detector). Figure (1.3) is an example of a TDR test of a flat 3-ohm cable that is damaged by wrapping a 1-inch piece of aluminum foil around the insulated cable. While no electrical contact is made, this minor perturbation of the transmission line properties can be detected by ECATS. The display shows the time-domain waveform with and without the foil wrap. The horizontal axis represents time in nanoseconds and the vertical axis represents voltage. The shape of the initial monocycle wavelet is indicative of the wave shape of the transmitted interrogation pulse. The large monocycle wavelet at the far right side is the reflection at the load end of the flat cable. The capacitive damage is clearly shown and its location on the cable is easily calculated.

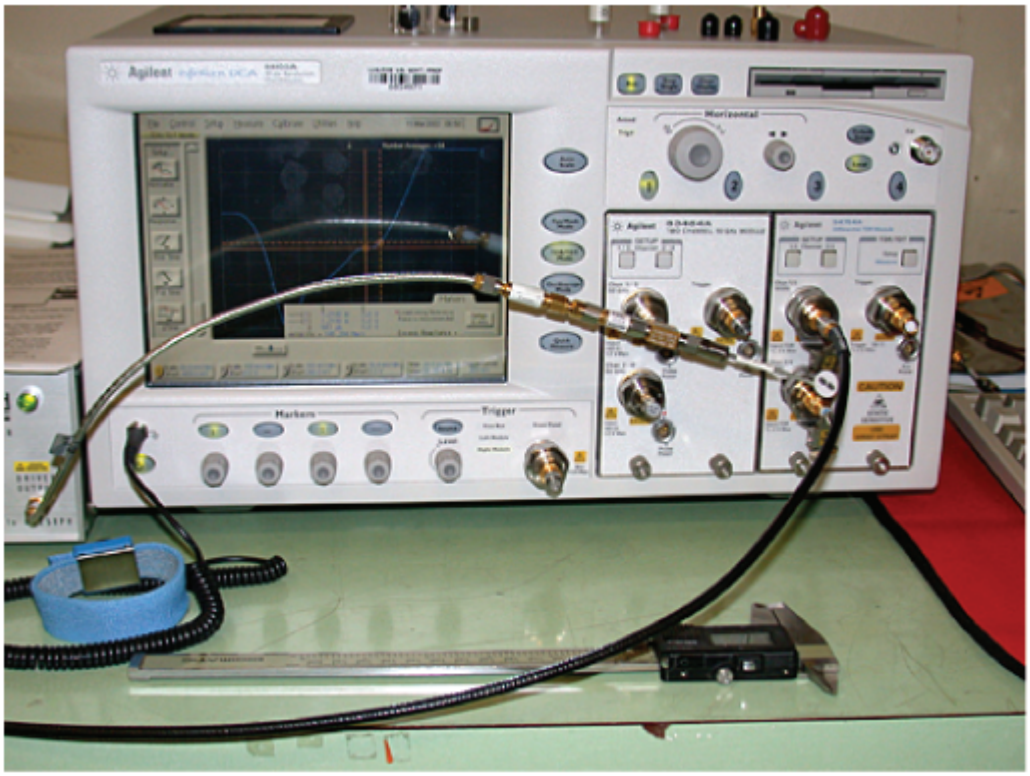


Figure 2.1: The Agilent Technologies TDR system used in ECATS is depicted in the figure.

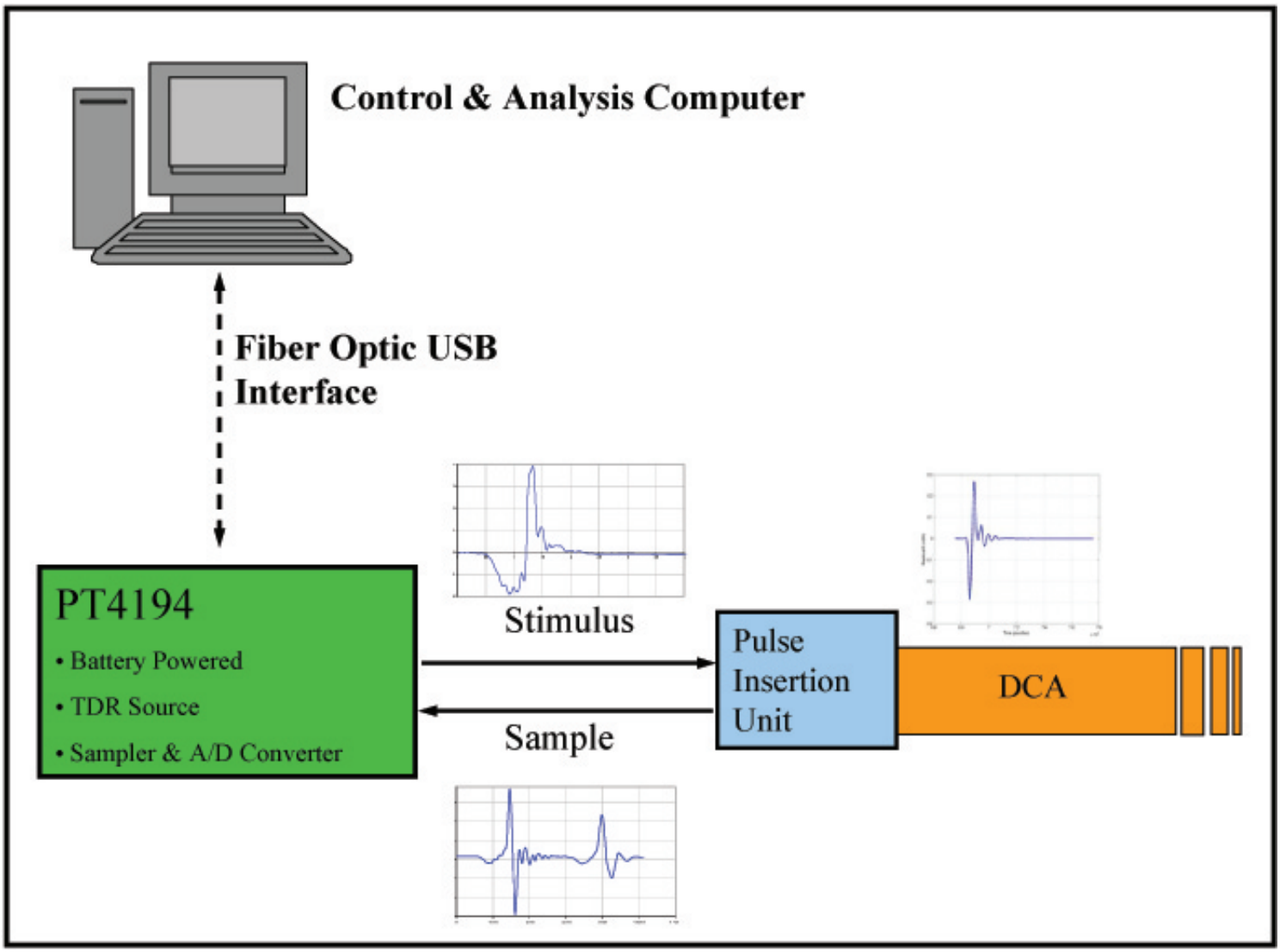


Figure 2.2: The ECATS measurement system block diagram is summarized in the figure.

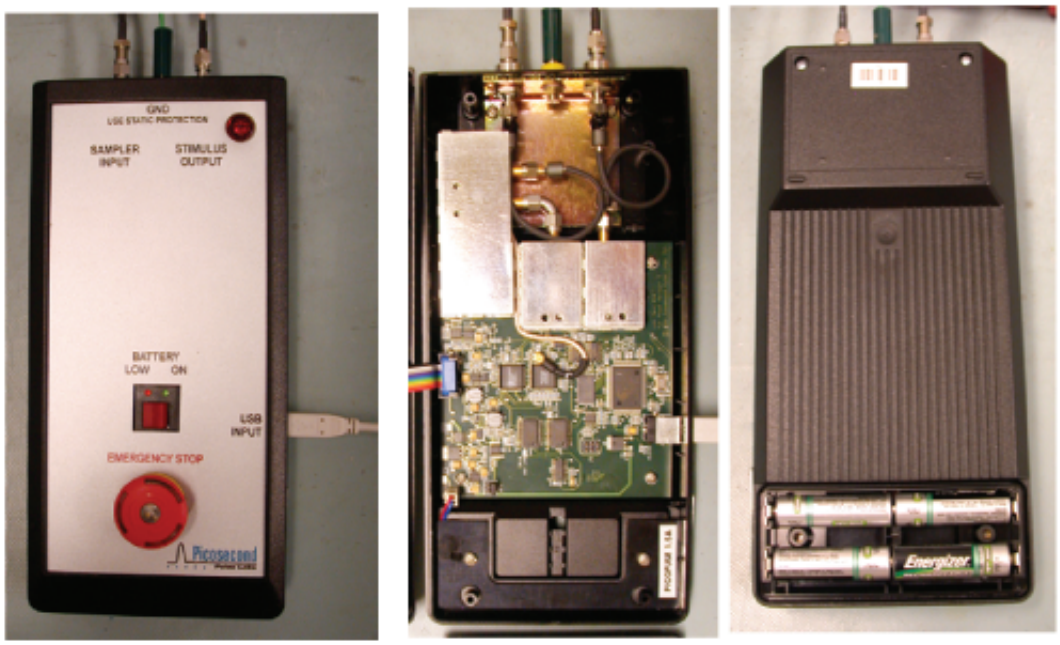


Figure 2.3: ECATS Figure 3: TDR Tester (PT4194) Pre-Production Unit is shown with red Emergency Stop plunger and with covers removed.

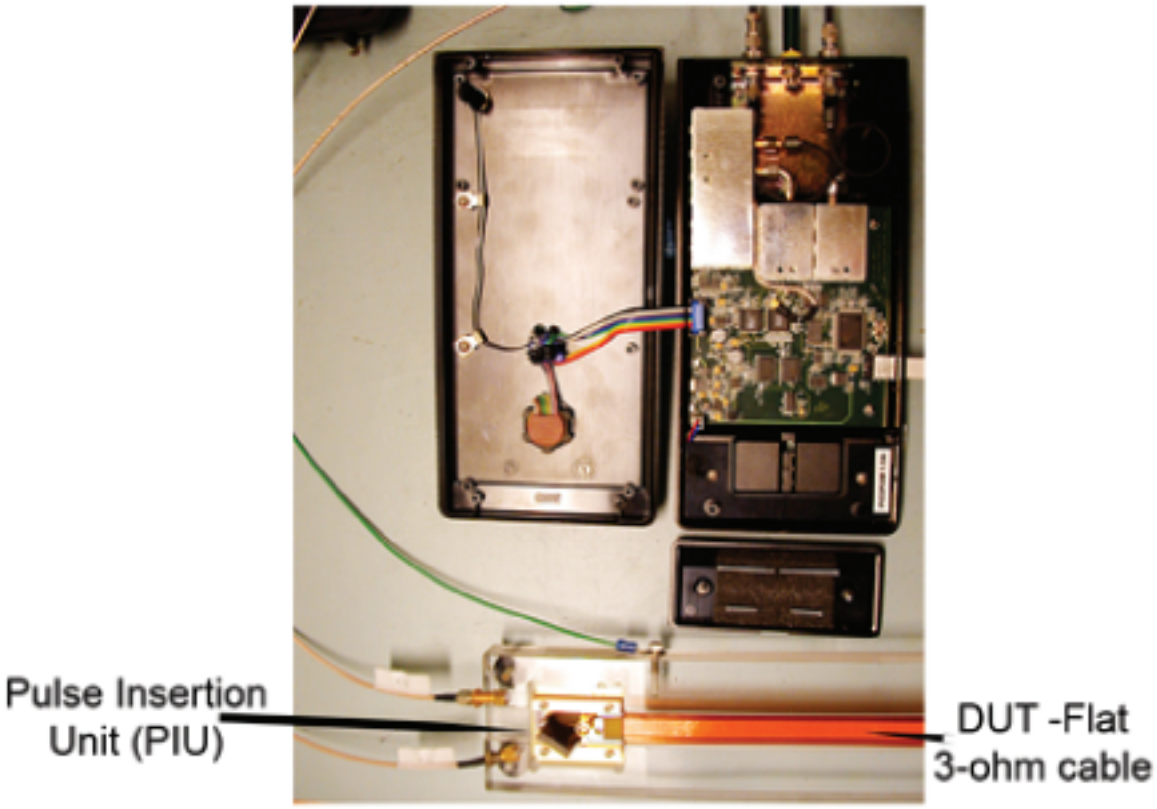


Figure 2.4: ECATS: The TDR Tester, the PIU and the DUT are depicted in the figure.

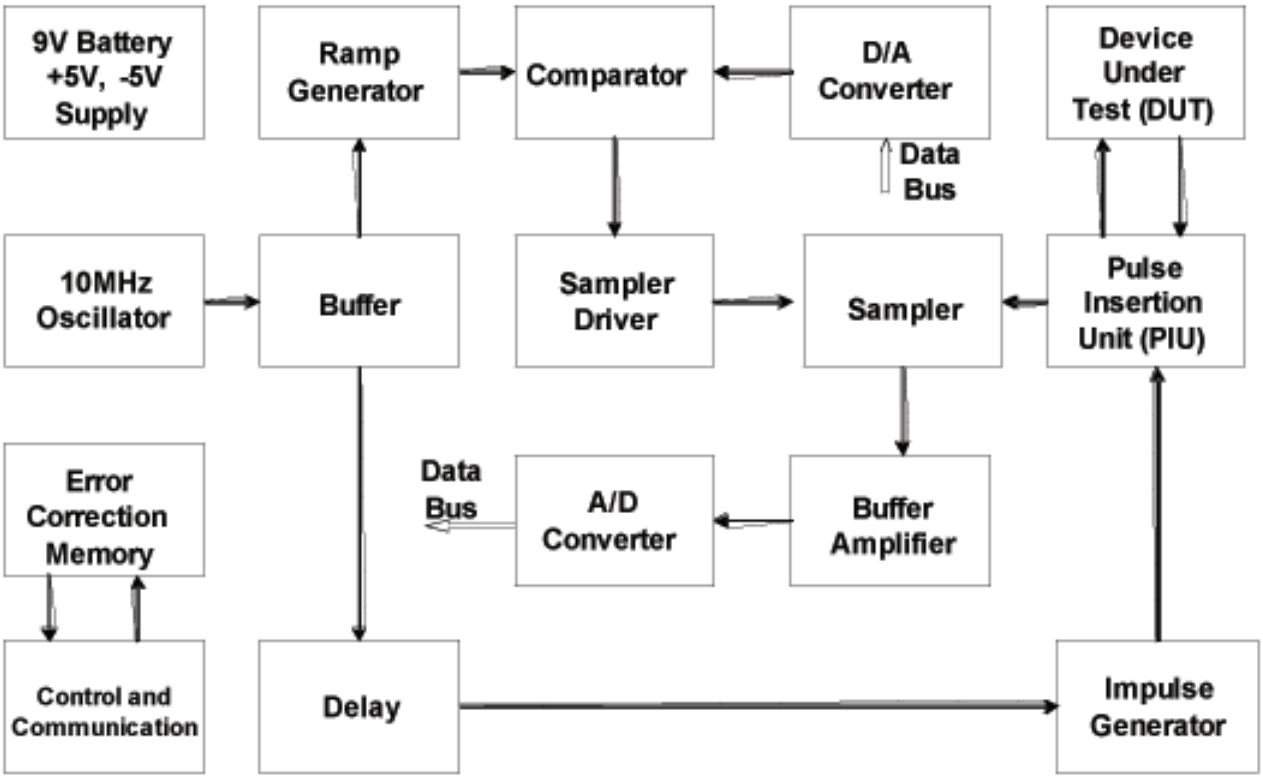


Figure 2.5: ECATS: The TDR Pulser Block Diagram is depicted in the figure.

2.4 Operational Safety

One of the most important aspects of ECATS design is its emphasis on safety. Operational safety was designed into the TDR tester from the beginning to assure that attachment to very sensitive circuits will not cause damage or induce a future failure mode. As shown in Figure (2.5), the design addresses the following safety issues.

2.5 General Protection

1. The TDR tester cannot output DC and limits output current to less than 10mA rms.

2. The only power source is four 1.2 volt AA batteries enclosed separately and fused.

3. Current-limiting resistors and DC-blocking capacitors are implemented in an independent, dual-series configuration. Dual-series resistors are each manufactured from different processing technologies.

4. The sampler board, sampler driver, stimulus, and current limit/test circuitry are each enclosed in separate grounded compartments for further protection.

2.6 Source Output Protection

1. In the tester, the stimulus impulse is generated on a short-circuited transmission line blocking the DC component to the pulse insertion unit (PIU).

2. The PIU capacitively-couples ($\sim$ 20 pF) the stimulus to the cable under test, blocking any DC component (redundant to Source Output Protection No. 1).

3. If both DC Blocks were to fail, DC current is limited (using dual-series resistors) to a maximum of 10mA.

2.7 Sampler Output Protection

1. DC is blocked from the sampler circuit strobe generator to the sample input by generating the strobe pulses on a shorted transmission line.

2. Dual series capacitors block DC from the sampler circuit strobe generator to the sample input (redundant to Sampler Output Protection No. 1).

3. Power supply currents of the sampler board are limited to less than 10mA using dual redundant current limiters. The limiters are verified before each test.

4. Current is limited to less than 10mA in case of A/D converter failure using dual-series current limiting resistors.

2.8 Tailoring Pulse Insertion Units to Cable Configurations

The modular design of ECATS allows testing of various transmission line geometries and impedances by simply changing the PIU. PIUs need to optimally match the geometry and impedance of the DUT, whether it is a flat or coaxial cable, 5-ohm or 50-ohm impedance. The most common PIUs could be kept in stock; for more exotic applications, a specialized PIU may need to be specially produced. Optimally matching the DUT with the PIU

reduces unwanted reflections, maximizes signal transfer to the DUT, limits the energy level coupled to the sensitive DUT, and enhances the signal-to-noise ratio, thus increasing measurement sensitivity (detecting smaller defects with less ambiguity).

2.9 Control and Analysis Software

The control and analysis software resides on a remote laptop computer, which is connected to the TDR tester through a standard fiber optic cable. The laptop can be located up to 30 meters from the tester and communicates via commercially available USB interface protocols. For safety reasons, this approach electrically isolates the TDR tester and PIU/DUT from the outside world, where power surges or lightning strikes may occur. The remote control software starts the TDR test. Prior to initiation of tester energy output, the control software verifies the TDR tester battery power and current limiter integrity. The resulting TDR waveform is stored in an ASCII test file. After waveform acquisition, the control software runs the analysis software. Once the analysis is completed, the control software displays whether the DUT was accepted or rejected. As a further safety precaution, the DUT is energized for the minimum required duration (much less than one second).

On site operators attach the PIU to the DUT and initiating the test at the remote computer. The control software evaluates the waveform to determine acceptance or rejection. All the TDR waveforms and control information are archived for reference. During preventative maintenance schedules, one can measure degradation and verify a mean-time-between-failure (MTBF). One might find the systems start failing soon after being exposed to a particular environment (e.g., high vibration, temperature, or humidity) or a particular use. Because ECATS provides quantitative measurements, customers can take advantage of these life-cycle insights.

2.10 Harware Competitors

The ECATS system developed by LLNL and Picosecond Pulse Laboratories was compared to existing products. The hardware from the following manufacturers was considered.

3M, Advanced Systems Tester, 900AST

AEA Technology, 20/20 TDR

Agilent, Digital Sampling Oscilloscope with TDR plug-in, 86100C and 545754

CM Technologies, Time Domain Reflectometer Card, PCI-3127

Eclypse International, ESP Plus

Intelligent Automation, WireCheck

Jovial Test Equipment, Shortstop

LeCroy, WaveExpert with TDR plug-in, SGA100 and ST-20

Northrop Grumman, AMWIT 1000

Phoenix Aviation, ARCMAS

Tektronix, Digital Serial Analyzer with TDR plug-in, DSA8200 and 80E04

See Figure (2.6) for a specification comparison of the ECATS system to existing products.

ECATS is the only cable testing device that can be safely attached to cables embedded or connected to explosive materials or other sensitive environments, and provide critical information about the electrical condition along the

Characteristic	LLNL & Picosecond Pulse Labs	Agilent [1] Scope + TDR [2]	3M [3]	Eclypse	Fluke
Model	ECATS	86100C + 54754A	900AST	ESP Plus	189
Technology	TDR	TDR	TDR	SWR [4]	Multimeter
Test Signal	Monocycle	Step	Step	Swept Frequency	DC
Amplitude of test signal at DUT [5]	150mV	200mV	60V	2Vp-p	1.2V
Duration of signal	800ps	2µs	3ns	CW	DC
Resolution of distance to short	<1mm	~1mm	18 inches [6]	18 inches [6]	N/A
Detects short/open	Yes	Yes	Yes	Yes	Yes
Detects degraded dielectric	Yes	Limited	No	No	No
Designed for low impedance cables	Yes	No	No	No	N/A
Battery Operated	Yes	No	Yes	Yes	Yes
Safe to use with explosive devices	Yes	No	No	No	No
Safety Features	Transformer coupling Back short at source 10mA current limits Fiber optic link to host	None	None	None	None
Portability	Handheld	Lab Instrument	Handheld	Handheld	Handheld
Cost	$25,000	$33,000	$4,800	$5,500	$430

Notes:

[1] Similar products are available from LeCroy and Tektronix

[2] Time Domain Reflectometry

[3] Similar products are available from AEA Technology, CM Technologies, Northrop Grumman, and Phoenix Aviation

[4] Standing Wave Reflectometry

[5] Device Under Test

[6] Tests performed at Sandia National Laboratory

Figure 2.6: ECATS: Comparison of ECATS with competing systems.

length of the cable under test. There are no other cable-testing devices that can measure and evaluate the conditions of cables to the degree of accuracy necessary for failure analysis throughout the life-cycle of these cables, from cable fabrication to final disposition (birth to burial). ECATS provides a truly unique tool for evaluating the integrity of low-impedance cables that are attached to explosive devices. It provides superior spatial resolution and improved detection of degraded or defective cables. A remote computer controls the data acquisition process by way of a fiber optic link. Dangerous voltages and currents are isolated from the data acquisition system.

The Agilent TDR (see Figure (2.1) makes measurements similar to ECATS; however, it is primarily a laboratory instrument and cannot be safely attached to cables embedded in sensitive environments, for which ECATS is made. Devices like the Fluke Multimeter (volt-ohm meter) can be approved for sensitive-cable measurements; however, they measure only cable continuity.

2.11 Principal Applications and Other Applications

The principal application of ECATS is to quantitatively measure, record, and automatically determine a Pass or Fail condition of a transmission line embedded in a sensitive or hostile environment, and if necessary, to periodically repeat this operation for long-term asset failure analysis. ECATS has been design to be extremely safe to use and operate, produce test results that are repeatable to a high degree of certainty and that can withstand rigorous scrutiny. Even after multiple hardware faults, the design precludes unwanted energy being transmitted to the DUT as these failures could be costly and possibly even life-threatening.

ECATS quantitatively evaluates and characterizes transmission lines, especially lines embedded in various systems. Transmission lines can have various geometric configurations, such as flat or co-axial, and can have various characteristic impedances, ranging from a few ohms to hundreds of ohms. ECATS can be used in long-term maintenance programs to establish failure rates and guarantee continued system reliability. Applications can also include quality control on production lines, from initial cable fabrication to lines embedded in subsystems and final system assembly. ECATS is also appropriate for military and commercial aircraft, satellites and missiles, with the possibility of embedding ECATS units in the high-value system for continued monitoring of critical cable-related functions. Another application is embedding ECATS units in systems under development that must withstand vigorous environmental testing, such as shock and vibration, temperature extremes, or submersion. In these applications, ECATS will continuously monitor and record cabling conditions for failure analysis. An important ECATS capability is determining exactly where along the transmission line stress or destructive forces are being applied to the transmission line.

Chapter 3

Test Fixtures and Measurements

3.1 Test Fixtures: 2D and 3D "Mockups "

The TDR waveform is very sensitive to the material surrounding the cable (especially metals that completely surround the cable) as well as the contour to which the cable conforms. Therefore it was necessary to design a test fixture that would simulate the cable's physical environment in a weapon. In addition, the TDR test fixture should allow for easy installation and removal of cables in order to test many cables both before and after introducing damage. These requirements resulted in the design of the two-dimensional (2D) TDR test fixture which was used for most of the test data collected thus far. In addition, 46 cable assemblies that had been removed from weapons were were tested in this fixture. The results of these tests were for information only and were used for repeatability and analysis development. See SIER 20062947LL and the instrument safety committee HIEC memo WESG-2006-06-01. The part number for the TDR 2D Test Fixture is AAA05-508170.

While the 2D test fixture will produce waveforms very similar to what is expected in a real weapon, there will still be some differences as the 2D fixture design contains supplemental features to facilitate cable installation and removal. Additionally, surrogates were used for certain materials that could result minor differences. At this time is not possible to capture TDR data on a real weapon due the extensive safety analysis, documentation, and review that has not yet been finalized or completed. Therefore, in an attempt to capture a gold standard waveform to simulate a TDR signal from a good cable in real weapon, a high fidelity 3D test fixture was created and tested. All tests with the 3D fixture utilize the same cable as the significant effort required to change out the cable makes doing so infeasible. The part number for the TDR three-dimensional (3D) Test Fixture is AAA06-53363-AA.

3.2 Measurement Bias, Truncation and Alignment

Figure (3.1) depicts an overlay plot of five representative signals that are measured for five different cables. It is readily seen that the signals contain various levels of zero-frequency (DC) bias. The signals were necessarily truncated before they were able to ring down to zero because the transient digitizer allows a fixed number of samples to be acquired (1024). The sampling period is chosen to satisfy the Nyquist sampling theorem and still allow the signal to be acquired. Later in the processing, the signals are truncated further for damage detection analysis (see the section on signal processing). Less obvious from this figure is the fact that the signals are not perfectly aligned in time (see the discussion below).

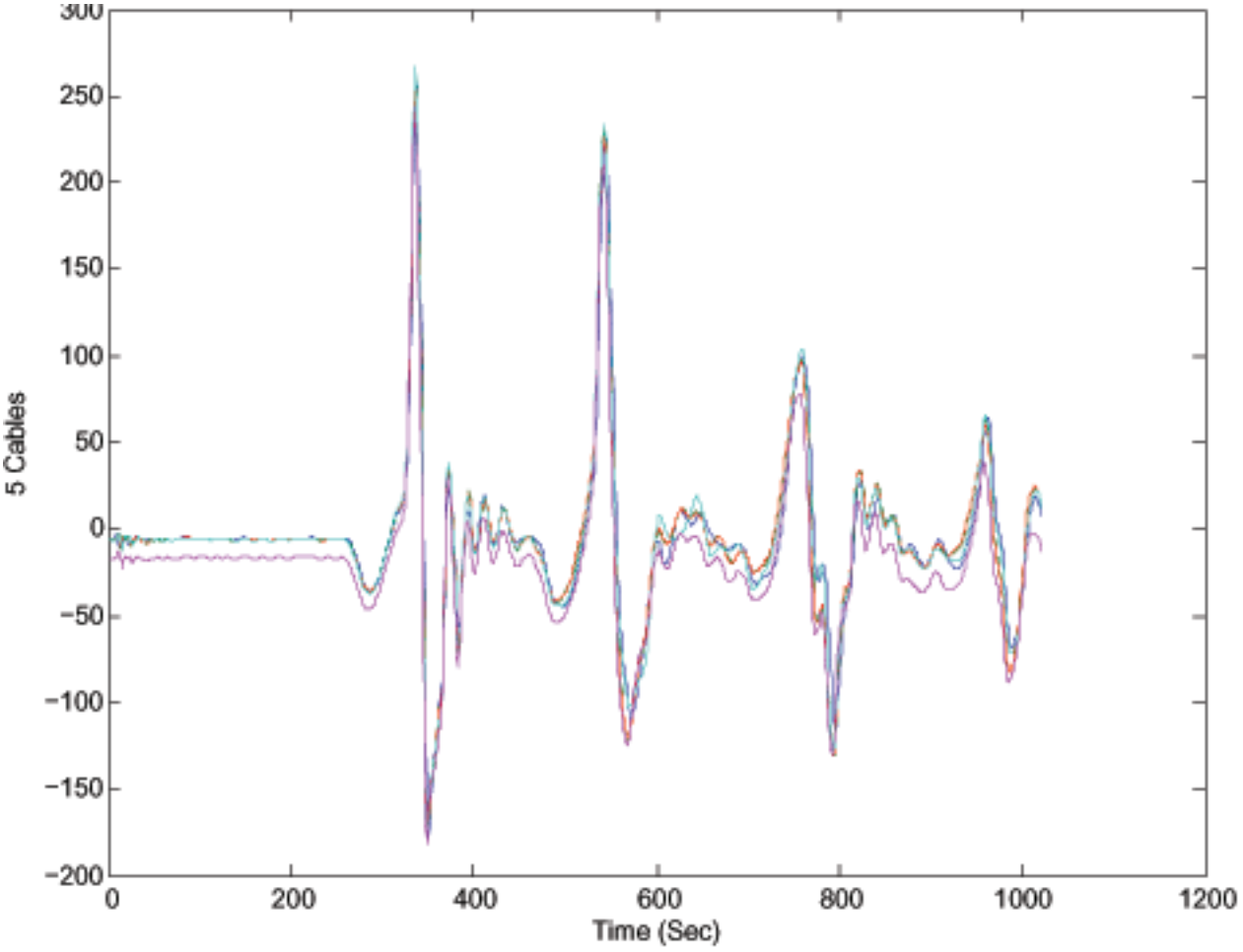


Figure 3.1: Raw TDR measurements for five different cables overlaid on the same plot. Note that they are biased and truncated in time. Less obviously, they have small variations in time delay. These are all factors with which we must cope using signal processing algorithms.

3.3 Automatic Signal Alignment Algorithms

It was found that the measured TDR signals contain a small, but significant variability in the start time of the pulses from pulse to pulse. This manifests as a small delay that appears among the TDR signals. Because we are doing comparisons with a standard pulse, it is important to remove these delays before computing innovations (residuals or error signals). The delays could create artificial pulses in the innovations and lead to false detections. Therefore, we developed automatic signal ensemble alignment algorithms to remove the delays.

Given an ensemble of $N + 1$ TDR signals, the alignment algorithms include the following basic steps:

• Selection of one signal from the ensemble to use as a reference signal to which the other signals will be compared. This leaves N signals for alignment.

• Mean and trend removal to eliminate any zero frequency bias in the signals.

• Interpolation (upsampling) by a large factor (usually ten) to allow a fine adjustment of delay. Both the reference signal and the signal under test are interpolated.

• Time delay estimation: The cross-correlation function between the reference signal and one signal from the ensemble is computed over a reasonable range of delays (usually about one-fourth to one-half the length of the signal). The largest peak in the cross-correlation is found. The delay at which the peak occurred is used as an estimate of the delay between the the signal under test and the reference signal. Note that the delays we can estimate are quantized in very small increments because we interpolated the signal by a factor or ten or so. This allows a very fine adjustment. This is sometimes referred to as a way to achieve "sub-sample resolution. "

• Time shifting: The signal under test is shifted (delayed or advanced) by the amount of the estimated delay. Care is taken to ensure that zeros are padded at the beginning or the end of the signal, as appropriate to keep the number of samples in the signal standard.

• Decimation (downsampling) to return the sampling period of the shifted signal to that of the original measured signals for later analysis.

• This procedure is applied to all of the N signals to be aligned until all of the signals have been aligned to the reference signal.

This automatic alignment procedure has reduced the variability in the TDR signals significantly and improved the ability to detect smaller amounts of damage. See Figures (3.2) and (3.3).

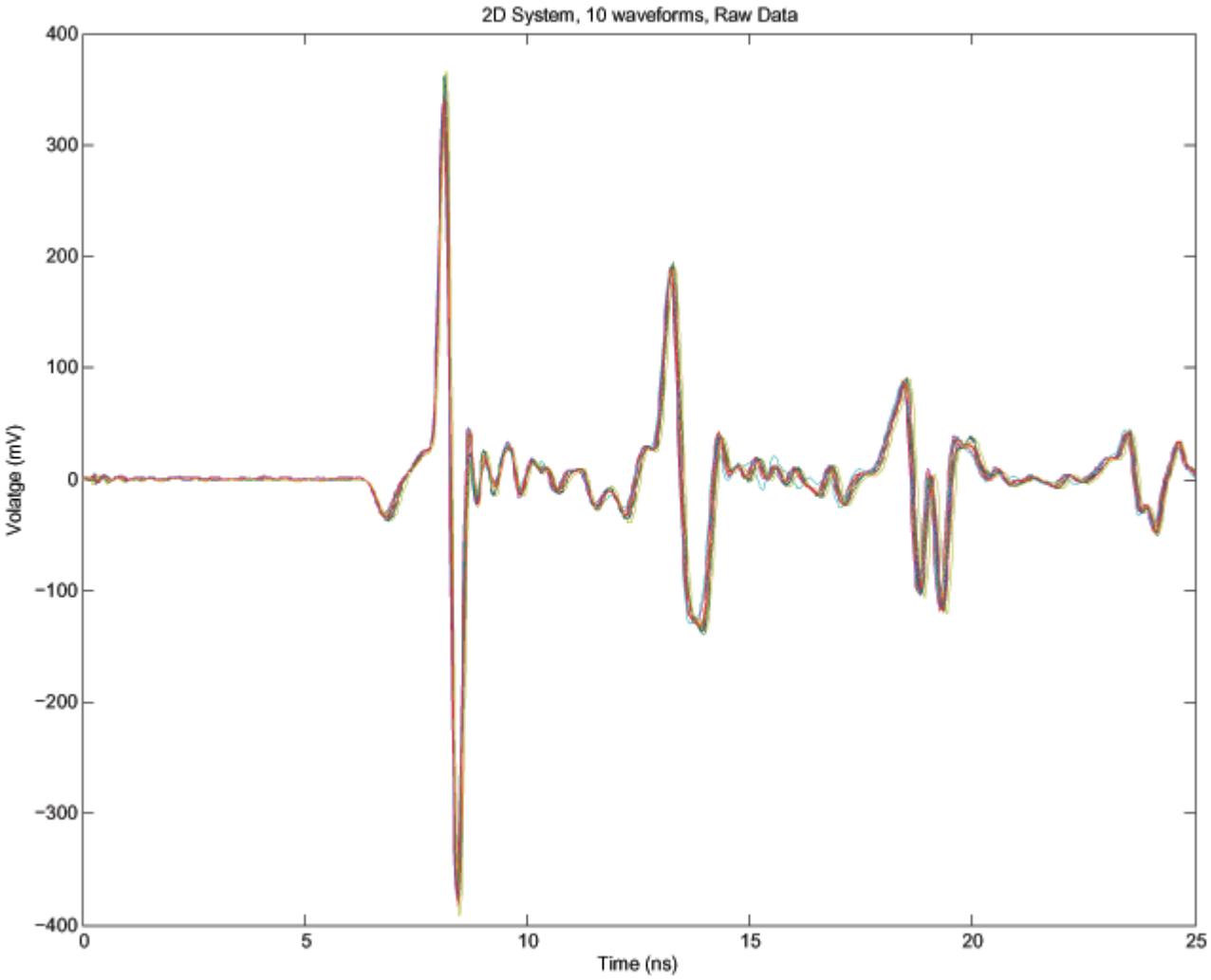


Figure 3.2: Before Pre-Processing with auto-alignment. The figure depicts an overlay plot of ten raw waveforms taken on the 2D Mockup fixture. No pre-processing or auto-alignment have been done at this step.

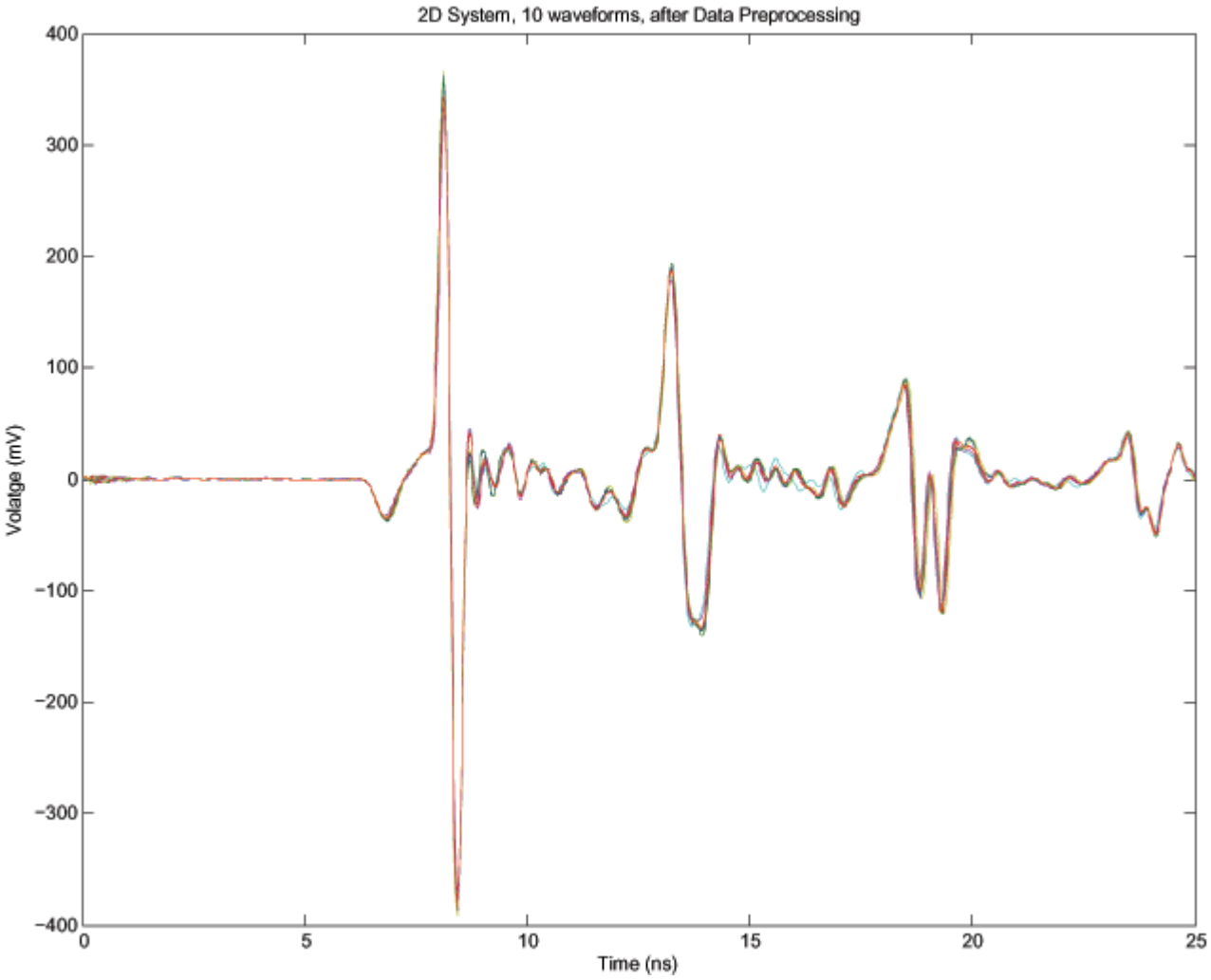


Figure 3.3: After Pre-Processing with auto-alignment. This figure depicts an overlay plot of the same ten waveforms taken on the 2D Mockup fixture, except that now they have been pre-processed and aligned by the auto-alignment algorithm described in this section. It is clear from visual inspection that the variability in the signals has been dramatically reduced.

Chapter 4

Measurement Repeatability

4.1 General Observations about Repeatability

In any comparative technique, the most important assumption is that one can reliably compare measurements from a standard with measurements from a device under test. Of course, this also assumes that the measurements are repeatable enough to allow such comparisons. These concepts are especially true for cable damage detection.

The measured TDR signals experience significant variability. For a given cable, the factors that affect variability include:

(1) The cable environment, including the materials around the cable and their proximity to the cable, (2) The attachment of the cable to the fixture, (3) The attachment of the cable to the Pulse Insertion Unit, (4) The measurement equipment, (5) Statistical Sample Size: We must live with the fact that the number of available cables is low (about 61) and the number of measurments we can acquire for each cable is low (about 3). The consequences are that the statistical confidence intervals about the probability of correct classification P_{CC} are wide. (6) Operator consistency - the technician doing the experiments learns techniques to make the measurements more consistent as she/he develops experience with the equipment. (7) The cable-to-cable variability for known undamaged cables.

This project has led to the following specific conclusions regarding repeatability:

(a) Once the cable is attached (inserted into the fixture and connected to the pulse insertion unit (PIU)), the TDR signal has very little variability from measurement to measurement.

(b) In the 2D fixture, from attachment to attachment for a single cable, the variability is significant, but reasonably small.

(c) From cable to cable, the variability is even greater, and presents a significant issue. In fact, the variability from *cable to cable* is the greatest barrier to detection sensitivity (the "long pole in the tent ") for the damage detection algorithms. The detection algorithms are very sensitive to small differences between the reference standard signal(s) and the signal(s) under test. However, when the measurements are not repeatable, the detection algorithms must be tuned to be less sensitive, so they do not mistake the variability in the measurements for damage. This is demonstrated in the receiver operating characteristic (ROC) curves depicted in the later chapters showing detection results.

Recall the definition of probability of correct classification, P_{CC}. For the cable damage detection problem, we let H_0 represent the hypothesis that no damage exists, and let H_1 represent the hypothesis that damage does exist. Recall that the key conditional probabilities are $P(H_1|H_1) = P_D$, the probability of detection (or sensitivity), and $P(H_0|H_0) = P_{spec}$, the specificity. From the contingency tables (confusion matrices) shown in the chapter on

detection perfomance measurement algorithms, we know that $P_{spec} = 1 - P_{FA}$, where P_{FA} is the probability of false alarm. Exploiting these results, we derive the following expressions:

$$P_{CC} = P(H_1|H_1)P(H_1) + P(H_0|H_0)P(H_0) \tag{4.1}$$

$$= P_D P(H_1) + [1 - P_{FA}]P(H_0) \qquad \text{(Probability of Correct Classification)} \tag{4.2}$$

If the application does not allow knowledge of the prior probabilities, a reasonable way to proceed is to assume that the two hypotheses are equally likely, so $P(H_1) = P(H_0) = 1/2$. Under this condition, the expression for P_{CC} is simplified to the following:

$$P_{CC} = \frac{1}{2}\{P_D + [1 - P_{FA}]\} \qquad \text{(Probability of Correct Classification)} \tag{4.3}$$

If the repeatability is high, the probability of correct classification is correspondingly high. The detection algorithm can be tuned to have a low probability of false alarm and a high probability of detection.

If the repeatability is low, the probability of correct classification is correspondingly low. This is because the detection algorithm must be tuned to allow a higher probability of false alarm, with a correspondingly lower probability of detection.

For the cable damage detection problem, some information about the cables is known a priori, so some estimates of the prior probabilities of the two hypotheses could possibly be made. For the purposes of this report, that work was not done because of programmatic constraints on resources. The priors were assumed to be equal. A study of prior probabilities is proposed for future work.

4.2 Cable "TDR Birth Certificates"

In an ideal application, each TDR signal for each cable would be measured at the time the cable was installed in the device. This would provide a "TDR birth certificate" for each cable, against which one could compare the TDR signals measured at a later time. This would obviate the issues of cable-to-cable variability. For our application, without such "TDR birth certificates," we must live with the cable-to-cable variability, try to mitigate it and accept the resulting limits to probability of correct classification.

4.3 Statistical Confidence Intervals About the Ensemble Mean Signal

4.3.1 Introduction

We developed some algorithms and software tools to analyze the repeatability of the measured TDR signals. The basic approach is to use statistical confidence intervals about an ensemble mean signal. Given an ensemble of TDR signals, we can compute its ensemble mean signal as depicted in Figure (4.1). We can now compute a time-varying statistical confidence interval about the mean signal and plot the bounds over the plot of the mean signal as in Figure (4.2). Now, given a new ensemble containing signals not included in the ensemble used to compute the bounds, we can test whether or not they lie within the bounds to give us an indication of repeatability. The remainder of this section describes the theoretical basis for the confidence interval bounds and how we use them for repeatability analysis.

Suppose we wish to estimate values of a scalar parameter $\theta(\underline{z})$ which is a function of a random vector $\underline{z}$ where

$$\underline{z} = \left[z_1, z_2, \ldots, z_M\right]^T \qquad \text{Vector of Measurements} \tag{4.4}$$

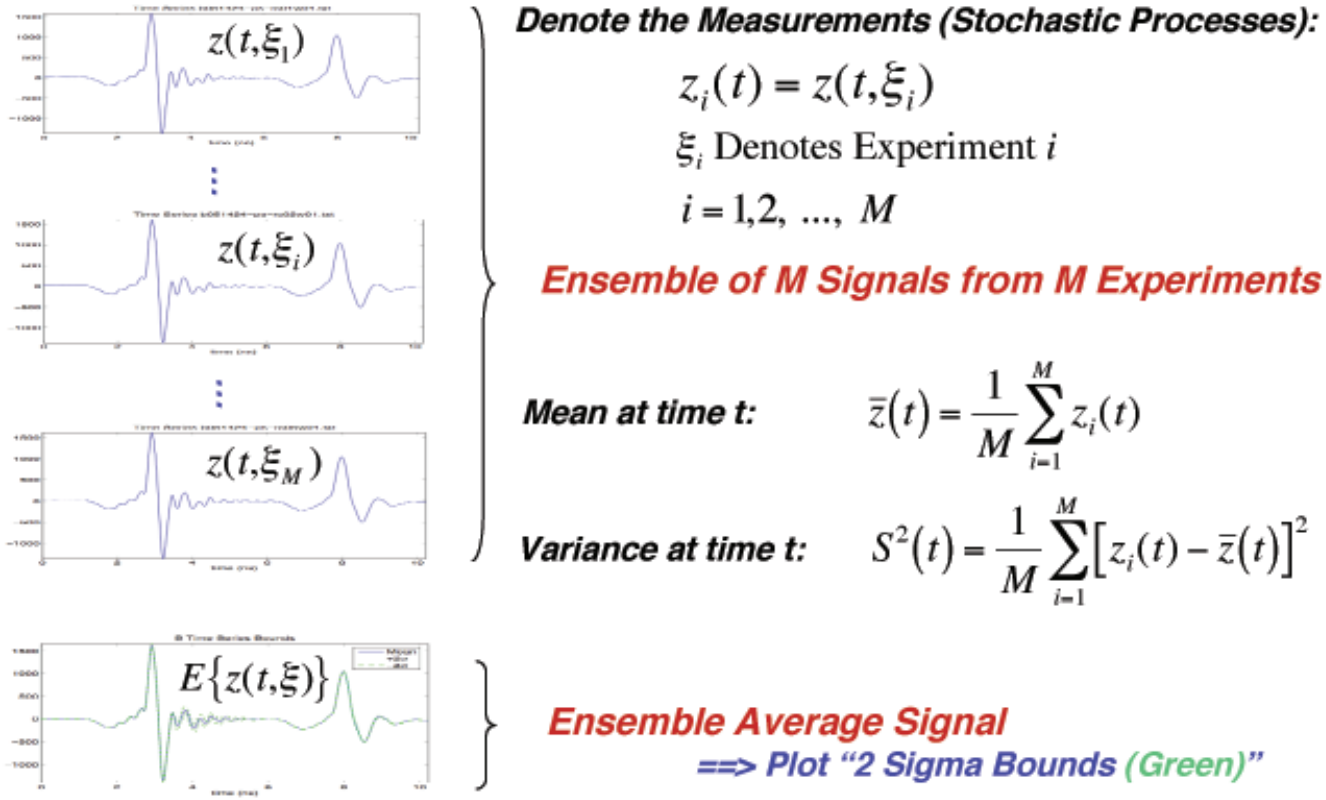


Figure 4.1: Ensemble averaging allows variability assessment. Let t = the time index and i = the experiment index.

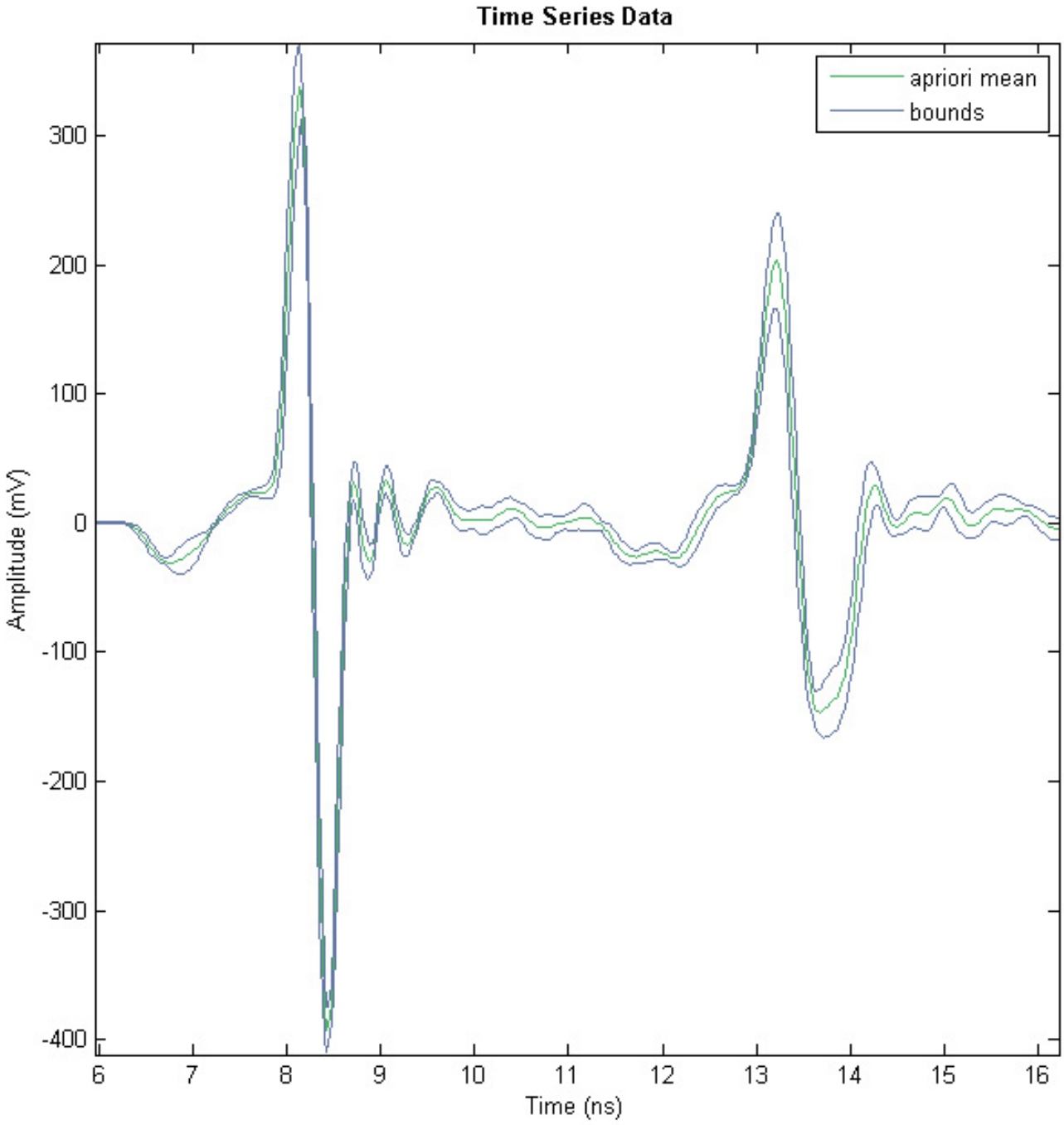


Figure 4.2: We plot the ensemble average (mean) signal and its confidence interval using the Student's t distribution formulation. This example is shown to demonstrate the principle of using confidence interval bounds because the bounds are not tight, so they are easily visible the reader. These data were acquired early in the project, when the repeatability of signals was not as good as it is currently.

and z_i is a random variable $i = 1, 2, \ldots M$. So, we have M measurements of a random variable. Consider a random interval (L, U) where L and U are functions of $\underline{z}$. Consider the probability that θ is contained in the interval (L, U):

$$P\{L < \theta < U\} = 1 - \alpha \qquad \text{Confidence Interval} \tag{4.5}$$

where α is a real number between zero and one and is called the significance of the test. We would like to determine the bounds L and U such that the above probability is true given a particular value of α. This discussion describes how we can derive the bounds (L, U) for the case in which θ is the mean of the random measurements and some particular assumptions are made. Most of the available theory assumes that a large number of measurements is available ($M > 30$), and this is called the large sample size case. For our problem, however, it is very difficult, if not impossible to gather signals in those quantities, so we must consider the small sample size case in which $M < 30$. [66, 49, 48, 67].

Suppose $z_1, z_2, \ldots, z_M$ are independent samples of a random variable. We define the sample mean as:

$$\bar{z} = \frac{1}{M} \sum_{i=1}^{M} z_i \qquad \text{Sample Mean} \tag{4.6}$$

The sample variance is defined as:

$$S^2 = \frac{1}{M} \sum_{i=1}^{M} (z_i - \bar{z})^2 \qquad \text{Sample Variance} \tag{4.7}$$

4.3.2 The Large Sample Size Case ($M > 30$)

Suppose that $z_1, z_2, \ldots, z_M$ are independent samples from a Gaussian distribution $N(\mu, \sigma^2)$ where the mean μ and the variance σ^2 are both unknown. We wish to estimate μ using the sample mean and calculate a confidence interval for μ. A confidence interval for this case can be derived to be the following [66]:

$$P\left\{\bar{z} - \frac{bS}{\sqrt{M}} < \mu < \bar{z} + \frac{bS}{\sqrt{M}}\right\} = 1 - \alpha \qquad \text{CI for Mean} \tag{4.8}$$

where b is a scalar constant found from tables of the normalized Gaussian random variable $N(0, 1)$ [66, 49, 48, 67]. As pointed out in [66], if σ is not known, the sample data will not yield a confidence interval with known end points. An alternate solution is to analyze the interval in terms of a student's t distribution, which is valid for both the small and large sample size cases. The case for which the random variable is not Gaussian is treated in [66]. For this case, we can compute only approximate intervals under limiting conditions.

4.3.3 The Small Sample Size Case ($M < 30$)

We wish to solve the problem posed earlier, except that $M < 30$. The following analysis is exact for both the small and large sample size cases [66]. To solve the small sample size case, we must define some more random variables and use the student's t distribution. When this is done, the confidence interval is given by

$$P\left\{\bar{z} - \frac{bS}{\sqrt{M-1}} < \mu < \bar{z} + \frac{bS}{\sqrt{M-1}}\right\} = 1 - \alpha \qquad \text{CI for Small M} \tag{4.9}$$

where b is a scalar constant found from a table of values of the student's t distribution. See Figure(4.3), Figure (4.4) and Figure (4.5).

- **For the Large Sample Size Case (M > 30):** *b = 1.96*
- **For the Small Sample Size Case (M < 30):**

M	b
2	6.31
3	2.92
4	2.35
5	2.13
6	2.02
7	1.94
8	1.90
9	1.86
10	1.83
11	1.81
⋮	⋮

Figure 4.3: For the 95% confidence interval, we use the values displayed for the factor b. When $M \geqslant 30$, use $b = 1.96$ from a table of the normalized Gaussian random variable. When $M < 30$, use the displayed constants for the factor b in Equation (4.9).

4.3.4 Comparison of the Small and Large Sample Size Cases

The table in Figure (4.6) shows how the small and large sample size bounds vary with the sample size M. Note that for Equation (4.8), the important factor is $1.96/\sqrt{M}$ and for Equation (4.9), it is $t_{.95}/\sqrt{M-1}$. Clearly, the interval bounds are monotonically decreasing in M, and

$$\frac{1.96}{\sqrt{M}} < \frac{t_{.95}}{\sqrt{M-1}} \qquad \text{for } M \geqslant 10 \tag{4.10}$$

The significance of this result is that for $M \geqslant 10$, the large sample bounds are always conservative with respect to the small sample bounds.

4.3.5 Confidence Interval for a Stochastic Process $z_i(n) = x_i(n) + v_i(n)$

In the previous sections, we were dealing with random variables. In our problem, however, we have stochastic processes in the form of our TDR signals. In the text, we denote the discrete time time index as $n = 1, 2, \ldots$. In the figures, we let discrete time be denoted by t. Figure (4.7) shows the notation for a stochastic process consisting of the sum of a deterministic (noiseless) signal and noise distributed $N(0, \sigma^2)$. Assuming that the deterministic signal and the stochastic noise are statistically independent, we obtain the relationships in Figure (4.8). We can now define the ensemble mean signal and the ensemble variance signal as depicted in Figure (4.9). For the large sample size case, we can now derive the $1 - \alpha$ confidence interval about the mean. The 95 percent confidence interval about the mean signal, as depicted in Figure (4.10). For the small sample size case, we can now derive the 95 percent confidence interval about the mean signal, as depicted in Figure (4.11).

4.4 Confidence Intervals Based on the Bootstrap

There exist some relatively new algorithms in the literature for coping with small sample size statistics in signal processing applications. These involve exploiting the concept of the "bootstrap" algorithm to compute the confi-

	alpha					
SampleSize	0.10	0.05	0.025	0.01	0.005	0.001
1.	3.078	6.314	12.706	31.821	63.657	318.313
2.	1.886	2.920	4.303	6.965	9.925	22.327
3.	1.638	2.353	3.182	4.541	5.841	10.215
4.	1.533	2.132	2.776	3.747	4.604	7.173
5.	1.476	2.015	2.571	3.365	4.032	5.893
6.	1.440	1.943	2.447	3.143	3.707	5.208
7.	1.415	1.895	2.365	2.998	3.499	4.782
8.	1.397	1.860	2.306	2.896	3.355	4.499
9.	1.383	1.833	2.262	2.821	3.250	4.296
10.	1.372	1.812	2.228	2.764	3.169	4.143
11.	1.363	1.796	2.201	2.718	3.106	4.024
12.	1.356	1.782	2.179	2.681	3.055	3.929
13.	1.350	1.771	2.160	2.650	3.012	3.852
14.	1.345	1.761	2.145	2.624	2.977	3.787
15.	1.341	1.753	2.131	2.602	2.947	3.733
16.	1.337	1.746	2.120	2.583	2.921	3.686
17.	1.333	1.740	2.110	2.567	2.898	3.646
18.	1.330	1.734	2.101	2.552	2.878	3.610
19.	1.328	1.729	2.093	2.539	2.861	3.579
20.	1.325	1.725	2.086	2.528	2.845	3.552
21.	1.323	1.721	2.080	2.518	2.831	3.527
22.	1.321	1.717	2.074	2.508	2.819	3.505
23.	1.319	1.714	2.069	2.500	2.807	3.485
24.	1.318	1.711	2.064	2.492	2.797	3.467
25.	1.316	1.708	2.060	2.485	2.787	3.450
26.	1.315	1.706	2.056	2.479	2.779	3.435
27.	1.314	1.703	2.052	2.473	2.771	3.421
28.	1.313	1.701	2.048	2.467	2.763	3.408
29.	1.311	1.699	2.045	2.462	2.756	3.396
30.	1.310	1.697	2.042	2.457	2.750	3.385
31.	1.309	1.696	2.040	2.453	2.744	3.375
32.	1.309	1.694	2.037	2.449	2.738	3.365
33.	1.308	1.692	2.035	2.445	2.733	3.356
34.	1.307	1.691	2.032	2.441	2.728	3.348
35.	1.306	1.690	2.030	2.438	2.724	3.340
36.	1.306	1.688	2.028	2.434	2.719	3.333
37.	1.305	1.687	2.026	2.431	2.715	3.326
38.	1.304	1.686	2.024	2.429	2.712	3.319
39.	1.304	1.685	2.023	2.426	2.708	3.313
40.	1.303	1.684	2.021	2.423	2.704	3.307
41.	1.303	1.683	2.020	2.421	2.701	3.301
42.	1.302	1.682	2.018	2.418	2.698	3.296
43.	1.302	1.681	2.017	2.416	2.695	3.291
44.	1.301	1.680	2.015	2.414	2.692	3.286
45.	1.301	1.679	2.014	2.412	2.690	3.281
46.	1.300	1.679	2.013	2.410	2.687	3.277
47.	1.300	1.678	2.012	2.408	2.685	3.273
48.	1.299	1.677	2.011	2.407	2.682	3.269
49.	1.299	1.677	2.010	2.405	2.680	3.265
50.	1.299	1.676	2.009	2.403	2.678	3.261
51.	1.298	1.675	2.008	2.402	2.676	3.258
52.	1.298	1.675	2.007	2.400	2.674	3.255
53.	1.298	1.674	2.006	2.399	2.672	3.251
54.	1.297	1.674	2.005	2.397	2.670	3.248

Figure 4.4: For the 95% confidence interval, we use the values displayed for the factor b. When $M \geqslant 30$, use $b = 1.96$ from a table of the normalized Gaussian random variable. When $M < 30$, use the displayed constants for the factor b in Equation (4.9).

	alpha					
SampleSize	0.10	0.05	0.025	0.01	0.005	0.001
55.	1.297	1.673	2.004	2.396	2.668	3.245
56.	1.297	1.673	2.003	2.395	2.667	3.242
57.	1.297	1.672	2.002	2.394	2.665	3.239
58.	1.296	1.672	2.002	2.392	2.663	3.237
59.	1.296	1.671	2.001	2.391	2.662	3.234
60.	1.296	1.671	2.000	2.390	2.660	3.232
61.	1.296	1.670	2.000	2.389	2.659	3.229
62.	1.295	1.670	1.999	2.388	2.657	3.227
63.	1.295	1.669	1.998	2.387	2.656	3.225
64.	1.295	1.669	1.998	2.386	2.655	3.223
65.	1.295	1.669	1.997	2.385	2.654	3.220
66.	1.295	1.668	1.997	2.384	2.652	3.218
67.	1.294	1.668	1.996	2.383	2.651	3.216
68.	1.294	1.668	1.995	2.382	2.650	3.214
69.	1.294	1.667	1.995	2.382	2.649	3.213
70.	1.294	1.667	1.994	2.381	2.648	3.211
71.	1.294	1.667	1.994	2.380	2.647	3.209
72.	1.293	1.666	1.993	2.379	2.646	3.207
73.	1.293	1.666	1.993	2.379	2.645	3.206
74.	1.293	1.666	1.993	2.378	2.644	3.204
75.	1.293	1.665	1.992	2.377	2.643	3.202
76.	1.293	1.665	1.992	2.376	2.642	3.201
77.	1.293	1.665	1.991	2.376	2.641	3.199
78.	1.292	1.665	1.991	2.375	2.640	3.198
79.	1.292	1.664	1.990	2.374	2.640	3.197
80.	1.292	1.664	1.990	2.374	2.639	3.195
81.	1.292	1.664	1.990	2.373	2.638	3.194
82.	1.292	1.664	1.989	2.373	2.637	3.193
83.	1.292	1.663	1.989	2.372	2.636	3.191
84.	1.292	1.663	1.989	2.372	2.636	3.190
85.	1.292	1.663	1.988	2.371	2.635	3.189
86.	1.291	1.663	1.988	2.370	2.634	3.188
87.	1.291	1.663	1.988	2.370	2.634	3.187
88.	1.291	1.662	1.987	2.369	2.633	3.185
89.	1.291	1.662	1.987	2.369	2.632	3.184
90.	1.291	1.662	1.987	2.368	2.632	3.183
91.	1.291	1.662	1.986	2.368	2.631	3.182
92.	1.291	1.662	1.986	2.368	2.630	3.181
93.	1.291	1.661	1.986	2.367	2.630	3.180
94.	1.291	1.661	1.986	2.367	2.629	3.179
95.	1.291	1.661	1.985	2.366	2.629	3.178
96.	1.290	1.661	1.985	2.366	2.628	3.177
97.	1.290	1.661	1.985	2.365	2.627	3.176
98.	1.290	1.661	1.984	2.365	2.627	3.175
99.	1.290	1.660	1.984	2.365	2.626	3.175
100.	1.290	1.660	1.984	2.364	2.626	3.174
inf	1.282	1.645	1.960	2.326	2.576	3.090

Figure 4.5: (Table Continued) For the 95% confidence interval, we use the values displayed for the factor b. When $M \geqslant 30$, use $b = 1.96$ from a table of the normalized Gaussian random variable. When $M < 30$, use the displayed constants for the factor b in Equation (4.9).

		Large Sample Case	Small Sample Case
M	$t_{.95}$	$\frac{1.96}{\sqrt{M}}$	$\frac{t_{.95}}{\sqrt{M-1}}$
2	6.31	1.385	6.31
3	2.92	1.13	2.06
4	2.35	.98	1.35
5	2.13	.879	1.06
6	2.02	.8	.90
7	1.94	.728	.792
8	1.90	.693	.706
9	1.86	.653	.657
10	1.83	.62	.61
11	1.81	.59	.573
⋮	⋮	⋮	⋮

Figure 4.6: Comparison of the Small and Large Sample Size Cases: For the 95% confidence interval, we use the values displayed for the factor b. When $M \geqslant 30$, use $b = 1.96$ from a table of the normalized Gaussian random variable. When $M < 30$, use the displayed constants for the factor b in Equation (4.9).

Denote the Measurements (Stochastic Processes):

$$z_i(t) = z(t,\xi_i) \ = Measurement$$
$$\xi_i = Experiment \ \mathrm{i}$$
$$i \ = 1,2,3,\cdots,\mathrm{M}$$

We Can Write the Measurements As the Sum of Signal Plus Noise:

$$i = 0,1,2,\cdots,M = \text{Experiment Index}$$
$$x_i(t) = \text{Noiseless Signal}$$
$$\mathrm{v}_i(\mathrm{t}) = \text{Noise distributed } \sim \mathrm{N}\left[0,\sigma^2\right]$$

$$z_i(t) = x_i(t) + v_i(t)$$

Figure 4.7: The stochastic measurement process is assumed to the sum of a deterministic signal and a stochastic noise signal.

$$z_i(t) = x_i(t) + v_i(t)$$

$$E\{z_i(t)\} = E\{x_i(t) + v_i(t)\}$$
$$= E\{x_i(t)\}$$

The Ensemble Variance Can Then Be Written:

$$\sigma^2_{z_i(t)}(t) = E\left\{\left[z_i(t) - E\{z_i(t)\}\right]^2\right\}$$
$$= E\left\{\left[z_i(t) - x_i(t)\right]^2\right\}$$
$$= E\left\{\left[x_i(t) + v_i(t) - x_i(t)\right]^2\right\}$$
$$= E\left\{v_i^2(t)\right\}$$
$$= \sigma^2(t)$$

Figure 4.8: Assuming the signal and noise signals are statistically independent, we can derive the variance waveform.

The test statistic is the ensemble mean

$$\mu(t) = E\{z(t)\}$$

Sample mean at time t (an estimate of the ensemble mean):

$$\bar{z}(t) = \frac{1}{M}\sum_{i=1}^{M} z_i(t)$$

Sample variance at time t:

$$S^2(t) = \frac{1}{M}\sum_{i=1}^{M}\left[z_i(t) - \bar{z}(t)\right]^2$$

Figure 4.9: We define the test statistic and estimate the sample mean and variance.

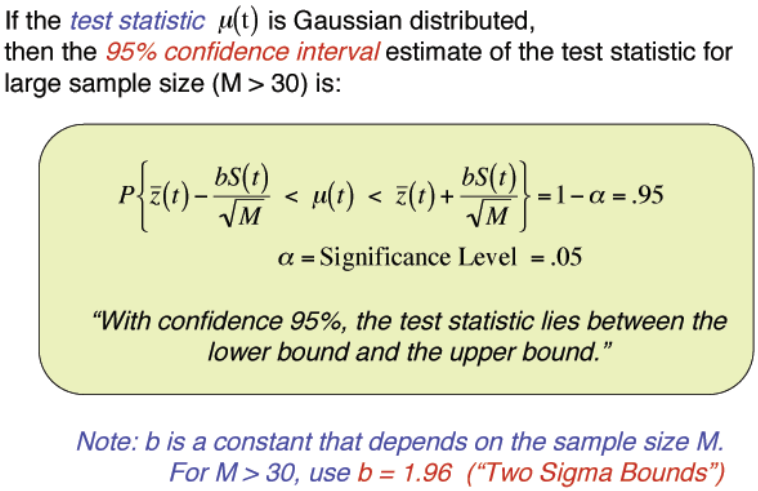


Figure 4.10: For the large sample size case, we compute a statistical confidence interval about the mean signal using the normal distribution.

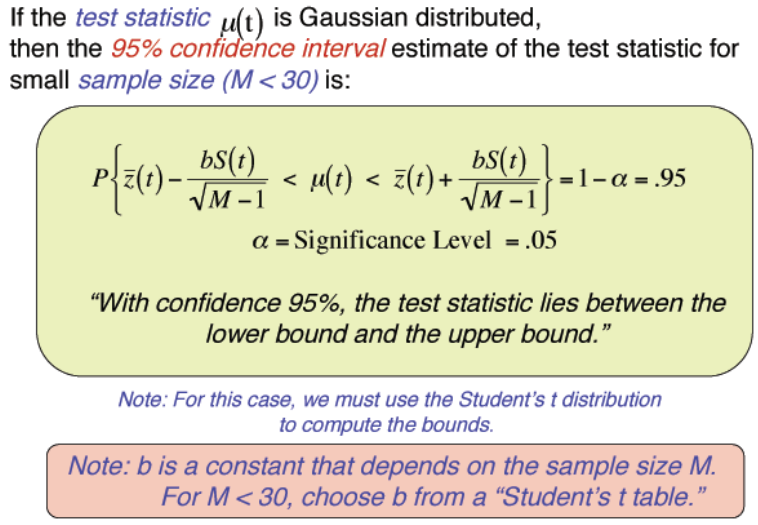


Figure 4.11: For small sample size, we compute a statistical confidence interval about the mean signal using the Student's t distribution.

dence interval on the sample mean when the sample size is small [24, 25, 26, 27]. We implemented an algorithm to compute the bootstrap mean and confidence interval on the mean. We included this as an option in the repeatability analysis code. Future work includes work with some advanced bootstrap algorithms that promise to have greater accuracy.

The bootstrap was introduced by Efron [26] as an approach to estimate confidence intervals for statistical parameters in circumstances in which standard methods cannot be applied. An example is when the statistical sample size is small because only a small number of data samples is available. The bootsrtrap does what the experimenter would do if it were possible - repeat the experiment many times to create a large sample size. The bootstrap approximates a large sample size by reassigning the observations randomly and recomputing key quantities many times (often thousands of times). These re-computations are then treated as repeated experiments.

The key to the bootstrap is very simple - resampling with replacement. Let $\mathcal{U}$ represent a finite population of individual units $U_1, U_2, \ldots, U_N$, any one of which is equally likely (with probability $1/N$) to be selected in a single random draw. A *random sample* of size n is defined to be a collection of n units $u_1, u_2, \ldots, u_n$ selected at random from $\mathcal{U}$.

In principle, the random sampling process uses a pseudo-random number generator to select independently a set of integers $j_1, j_2, \ldots, j_n$, each of which equals any value between 1 and N with probability $1/N$. These integers serve as indices that determine which members of $\mathcal{U}$ are selected in the random sample, $u_1 = U_{j1}, u_2 = U_{j2}, \ldots, u_n = U_{jn}$.

If *sampling with replacement* is used, then every sample is returned to the to the data set after sampling. This means that a single unit U_i is allowed to appear more than once, and some units may not appear at all. If *sampling without replacement* were used, then we would insist that the integers $j_1, j_2, \ldots, j_n$ be distinct. For the bootstrap, we use *sampling with replacement*.

For computing a bootstrap confidence interval on the mean TDR signal, we let $\mathcal{X}$ be the random sample vector of time samples in an ensemble of $x_U(n)$ signals used to compute an ensemble average. See Figure (4.1) for a depiction of an ensemble of signals used to form an ensemble average. We then use the following procedure to compute the bootstrap confidence interval on the ensemble mean signal:

Algorithm for Estimating the Bootstrap Confidence Interval About the Mean for a Random Sample Vector:

Step 0, Conduct the Experiment:
Let the Nx1 random sample vector be $\mathcal{X}$ as defined above.

Step 1, Resampling
Using a pseudo-random number generator, draw a random sample of N values with replacement from $\mathcal{X}$. We call this our bootstrap resample $\mathcal{X}^*$. Keep in mind that some of the original samples may appear more than once in $\mathcal{X}$ and others may appear not at all.

Step 2, Calculation of the Bootstrap Estimate of the Mean
Calculate a desired function of the of the values of $\mathcal{X}^*$. In our case, calculate the sample mean of the values in $\mathcal{X}^*$. Call this the bootstrap sample mean $\widehat{\mu}_1^*$.

Step 3, Repetition
Repeat Steps 1 and 2 a large number n of times (for example, let $n = 1000$) to obtain a total of n bootstrap esimtates of the the mean: $\widehat{\mu}_1^*, \widehat{\mu}_2^*, \ldots, \widehat{\mu}_n^*$.

Step 4, Approximation of the Distribution of $\widehat{\mu}$.
Sort the bootstrap estimates of the means into increasing order to obtain the vector $\widehat{\mu}_{(1)}^*, \widehat{\mu}_{(2)}^*, \ldots, \widehat{\mu}_{(n)}^*$, where $\widehat{\mu}_{(k)}^*$ is the kth smallest of $\widehat{\mu}_1^*, \widehat{\mu}_2^*, \ldots, \widehat{\mu}_n^*$. If desired, one can compute a probability density function (pdf) estimate for $\widehat{\mu}$ and plot it for visual inspection. This step is not necessary, but it is quite instructive. For example, one could calculate a histogram or a kernel estimate of the pdf.

Step 5, Compute the Confidence Interval

The desired $(1-\alpha)100\%$ bootstrap confidence interval is given by the set of lower and upper bounds $[L,U] = \left[\widehat{\mu}^*_{(q_1)}, \widehat{\mu}^*_{(q_2)}\right]$, where α is the significance of the test, the index on the lower bound is $q_1 = \lfloor N\alpha/2 \rfloor$, which is the integer part of $N\alpha/2$ rounded down, and the index on the upper bound is $q_2 = N - q_1 + 1$. These bounds were derived by finding the indices on the sorted means at lower and upper ends of the pdf estimate at which the areas under the pdf estimate equal α. This does not require that the estimated pdf is symmetric.

4.5 Confidence Intervals Based on Kernel Estimates of the pdf of a Signal Ensemble

We implemented a second alternative to using the confidence intervals defined earlier. The theoretical confidence intervals on the mean assume that the mean is Gaussian distributed. We can avoid this assumption by computing a kernel estimate of the pdf of an ensemble of reference $x_U(n)$ signals [64, **?**, 65]. Once this pdf estimate is known, it can be integrated to compute a confidence interval according a specified statistical significance α as for the theoretical case. Of course, the performance of kernel pdf estimators is compromised when the sample size is small. The repeatability code now contains an option test for repeatability using this kernel pdf estimate algorithm.

4.6 Repeatability Analysis Tool and GUI

We developed a repeatability analysis tool with a graphical user interface (GUI) in MATLAB. The tool performs several functions, including displaying plots of confidence interval bounds for TDR signals. Figures (4.12), (4.13),(4.14),(4.15), and (4.16) depict typical displays the user sees when using the tool. The tool allows the user to read in an ensemble of $x_U(n)$ signals, compute its ensemble mean signal, and plot the confidence interval bounds over the mean signal or the overlay plots of all of the signals in the ensemble. The confidence interval bounds can be computed in a variety of ways, as described in this chapter. These include (1) Using the theoretical bounds based upon the Student's t distribution, (2) Using "minmax" bounds, in which the upper bound is the maximum value of the ensemble signals at time n and the lower bound is the minimum value of the signals in the ensemble at time n. This is an ad hoc technique that gives some intuition, but does not provide theoretical rigor. (3) Using the estimated bounds on the sample mean based upon the bootstrap, and (4) Using the estimated bounds using a kernel estimate of the pdf of the ensemble signals.

Note that the GUI tool also plots a light blue "bar chart" analagous to a histogram of the number of signal samples that exceed the bounds at a given time n. This provides added information that contributes to our intuition. A region of interest, or region for analysis, can be defined using the red cursor lines depicted in the figures. The user can choose to compute the bounds, etc. only in this region. This is very convenient, because the early and late parts of the waveform are usually irrelevant to the test. The user can choose a variety of values for the significance of the test, α. Examples of the use of the repeatability GUI tool are displayed in Figures (4.12), (4.13), (4.14), (4.15), (4.16).

Repeatability GUI Help Package Description (Text from the Help Package in the Code)

This tool is used to create bounds from an a priori set of data that are then compared to a test set of data. The bounds created are based on the a priori data, the choosen significance 'alpha', and the choosen test type.

There are 4 different test types to compute bounds:

(1) minmax: All apriori data is bounded.

(2) kernel: A kernel estimate pdf at each time sample is computed and used to compute the bounds.

(3) student t: The theoretical bounds based upon the Student's t distribution.

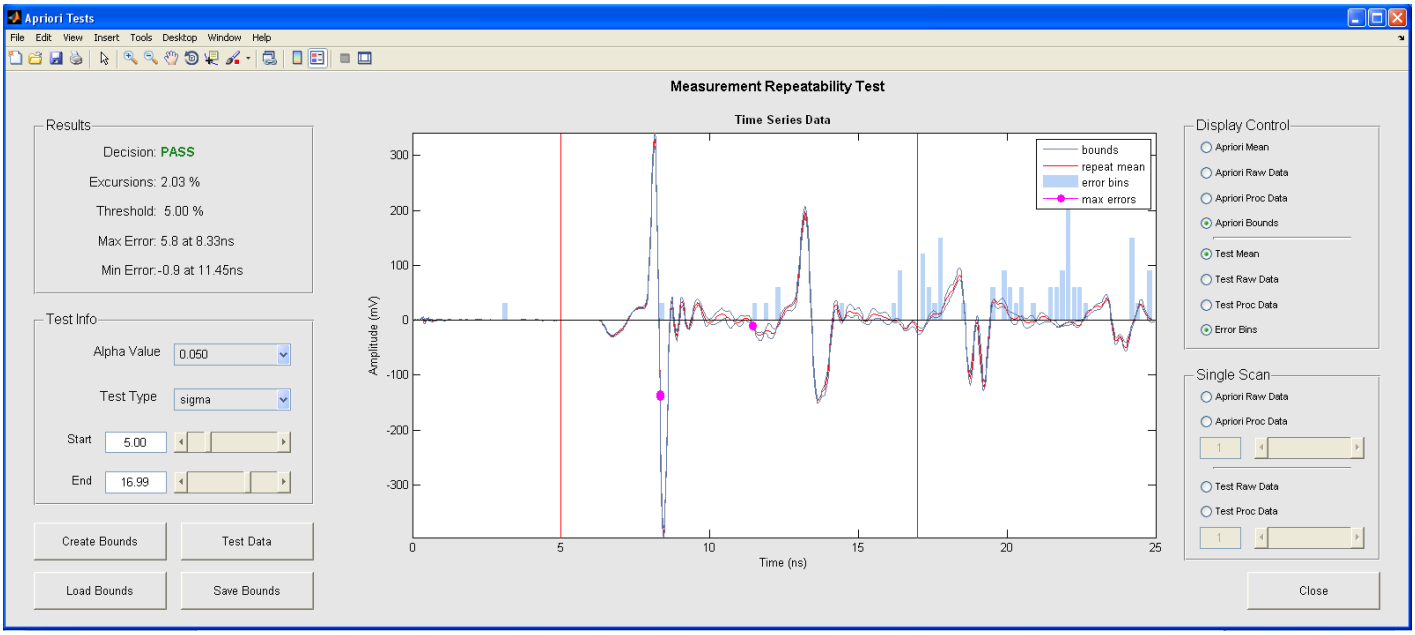


Figure 4.12: Student's t Bounds Test: Three signals for known undamaged cables were used to create the a sample mean and a priori bounds. Three signals from other known undamaged cables were used to create a sample mean signal, which was tested against these bounds. The cable set was declared to be repeatable. Note the vertical red lines depicting the region for analysis, and the light blue "histogram"bar chart representing the time samples at which the mean signal exceeds the bounds.

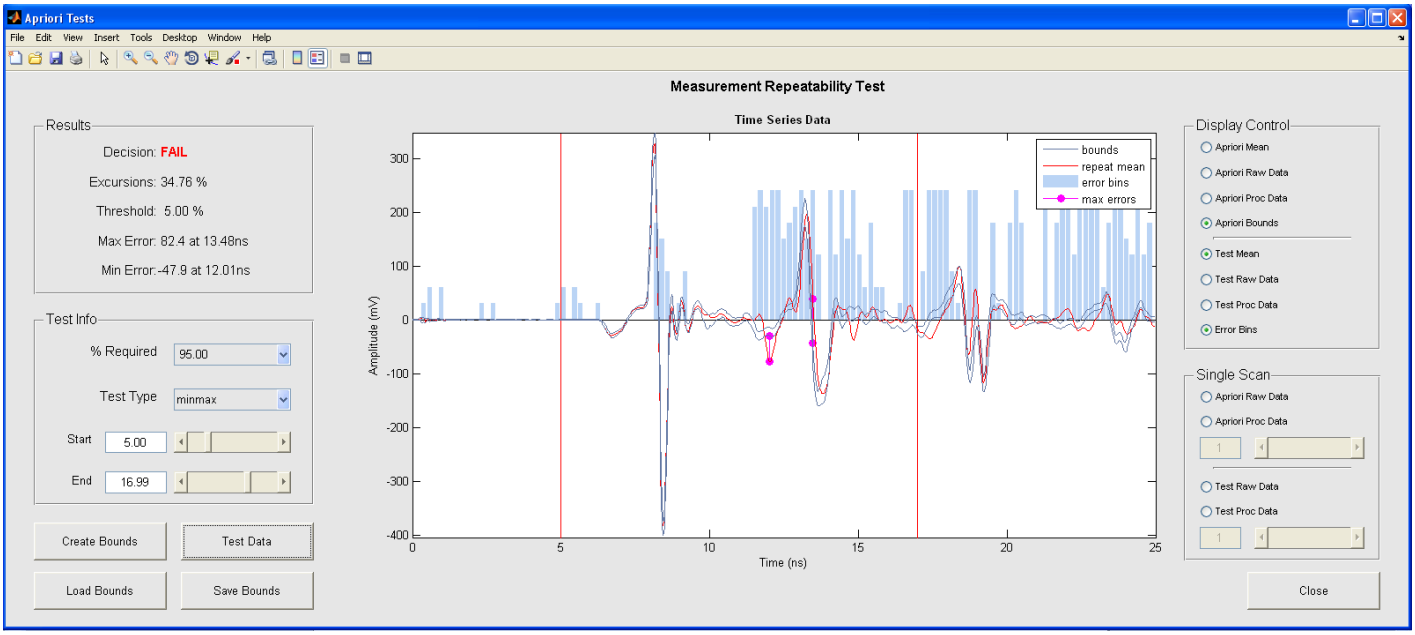


Figure 4.13: Minmax Test: Seventy-five signals for known undamaged cables were used to create the minmax a priori bounds. A single signal for a known damaged cable was compared against these bounds and was correctly indicated damaged.

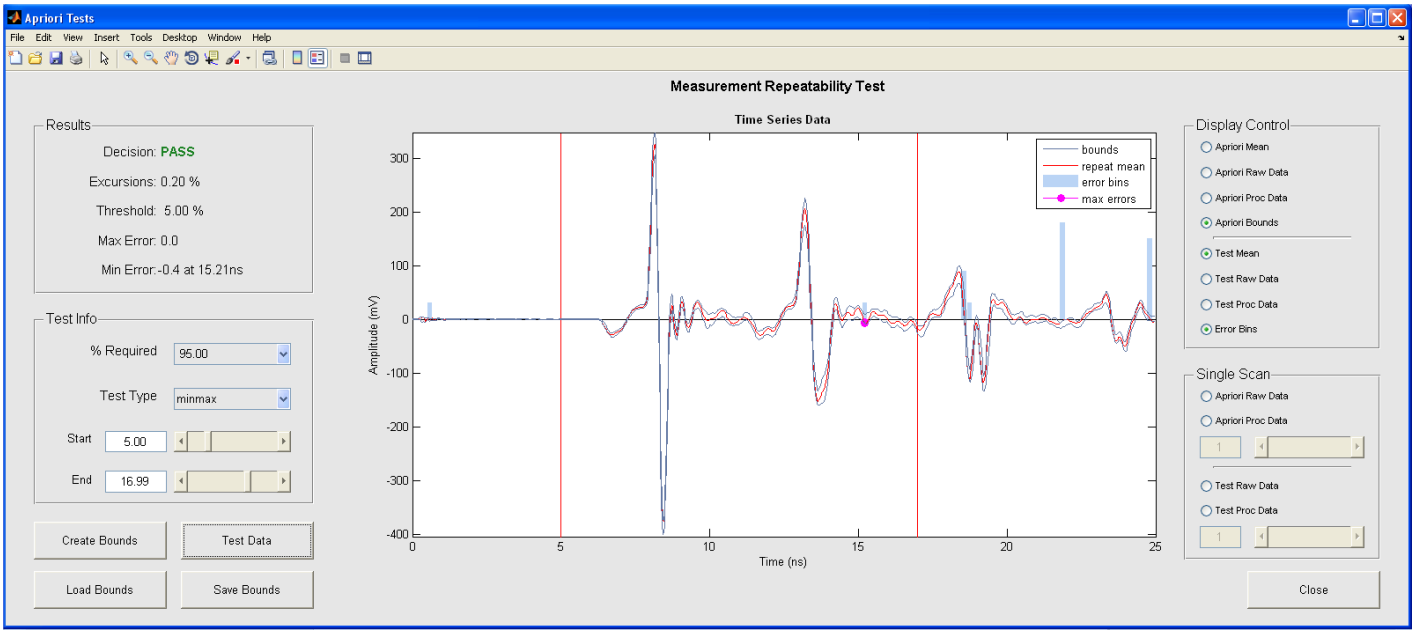


Figure 4.14: Minmax Test: Seventy-five signals for known undamaged cables were used to create the minmax a priori bounds. A single signal for a known undamaged cable was compared against these bounds and was correctly indicated undamaged.

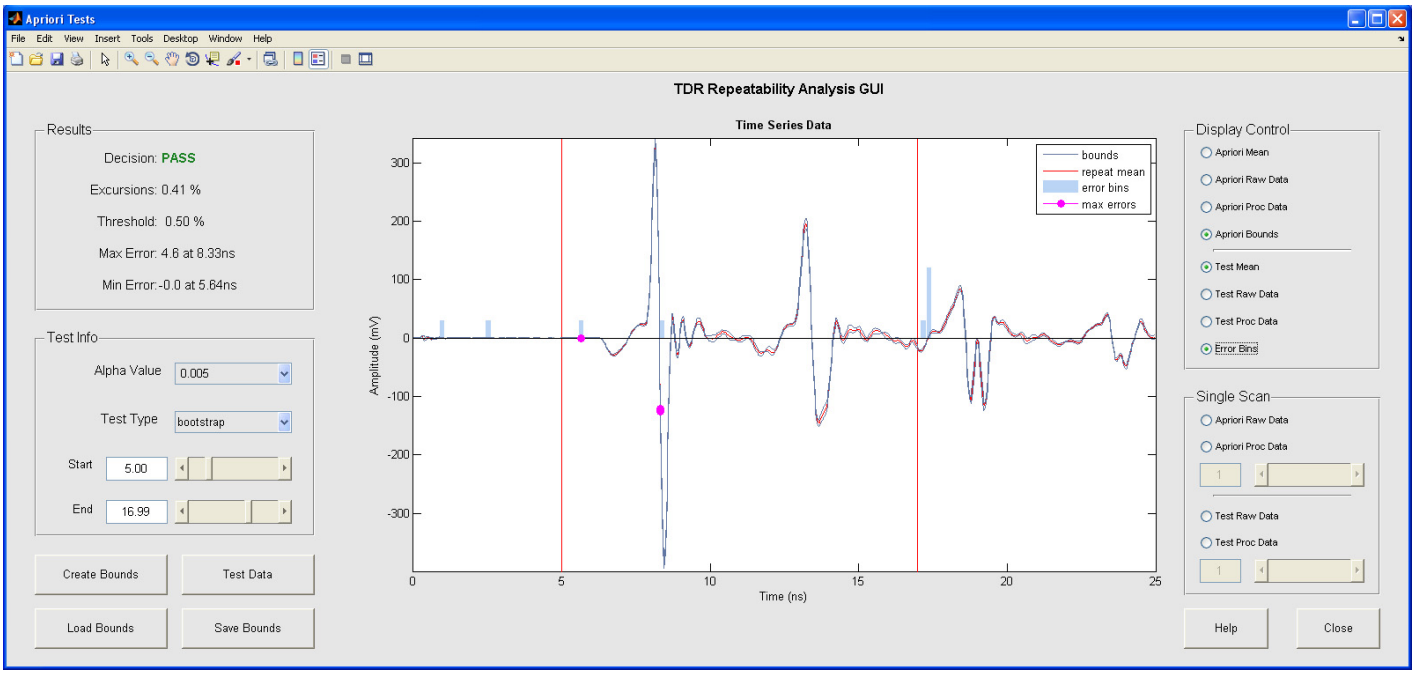


Figure 4.15: Bootstrap Confidence Interval: Five signals for known undamaged cables were used to create the bootstrap a priori bounds. Five signals for known undamaged cables were tested against these bounds and were correctly declared undamaged.

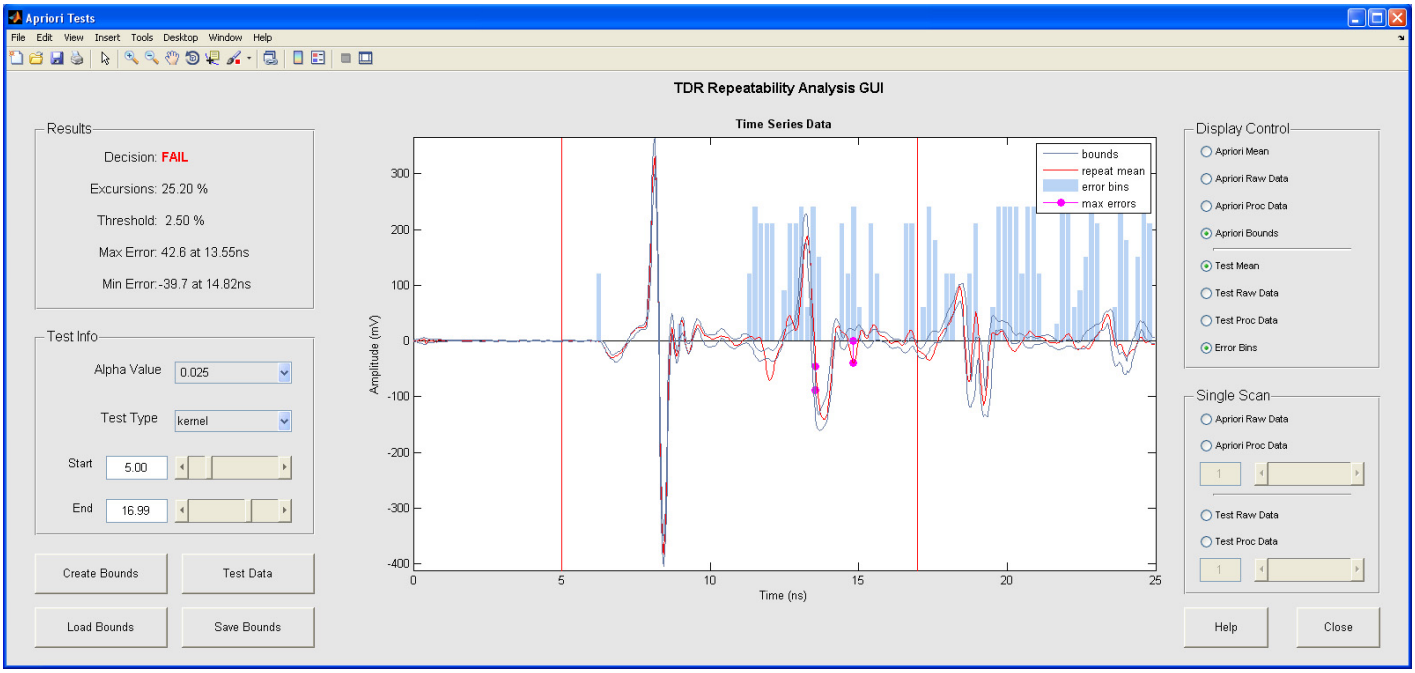


Figure 4.16: Confidence Intervals Based on Kernel pdf Estimation: Five signals for known undamaged cables were used to create the kernel-based a priori bounds. A single signal for a known damaged cable was tested against these bounds and was correctly declared damaged.

(4) bootstrap: The estimated bounds on the sample mean based upon the bootstrap re-sampled means.

The student t and minmax tests are useful for testing an ensemble of signals against the bounds for repeatability analysis. Both the a priori and test ensembles must be the same size for these tests.

The bootstrap and kernel pdf estimator algorithms are useful for testing both repeatability of an ensemble and for a single signal.

Only the region between the two red vertical lines shown on the plot are evaluated. This region can be changed by moving the start and end points in the Test Info pane.

Using the Display Control users can view the the a priori bounds, the mean from both the a priori and test set, and the full ensembles both raw and preprocessed (preprocessing demeans, detrends, and aligns). Individual waveforms can also be viewed by using controls in the Single Scan pane.

By selecting 'Error Bins' a light blue histogram displays the number of signal samples outside the bounds at a given time. The minimum and maximum deviations from the bounds are also shown.

4.7 Repeatability Studies and Results

4.7.1 Introduction

There are many variables associated with the data acquisition process which contribute to variation in the acquired data. In an effort to make quantitative comparisons among growing numbers of tests, it was necessary to take data sets that isolate each variable individually and to quantify the resulting effect. In addition to comparing the resulting data sets, it was also useful to examine means and confidence intervals of the various data sets to get a

better understanding of the repeatability. By comparing sets of tests in this way, it allows the tools and procedures to be honed to procedure a waveform with the best possible repeatability.

4.7.2 Repeatability Test Procedures

A key to producing repeatable TDR data was the development of a detailed procedure describing precisely how to setup and perform TDR tests. The use of a detailed procedure is also essential for the repeatability studies as any waveform variation must be attributed to the variable in question and not a consequence of the testing process. In addition to reducing the measurement-to-measurement repeatability issues, the procedure also helps mitigate operator-to-operator repeatability problems by providing a guide to assist in performing the data acquisition process reliably. The final procedure AF-PT4194 rev D is under configuration management in the Weapons Production & Quality Groups engineering authorization program Windchill, and is attached in Appendix A.

4.7.3 Static Repeatability

Initially it was believed the TDR tester produced repeatable waveforms if the test setup was not altered. This had been observed by the near perfect repeatability when collecting many waveforms, one after another. However, after seeing trends in the data related to when the data were taken, a test series was designed to examine the TDR tester's repeatability over time. In this series of tests, the pulse insertion unit (PIU) and cable assembly were not altered between tests and waveforms were collected every two minutes, therefore differences between waveforms can be attributed to two sources: variability caused by temperature changes within TDR electronic components, or to inherent repeatability limitations of the TDR (a consequence of its design). One would expect the latter to remain constant while the former would produce greater variability immediately after power-up and increased stability and repeatability after the electronics have had a chance to warm-up. Evidence of this phenomenon can be seen by comparing deviation from waveforms in Figure (4.17) to those in Figure (4.18). While both data sets have good repeatability, clearly there is an improvement after the tester has been powered-up for 30 minutes. When the set of waveforms was collected after the 30 minute warm-up, variations in waveforms are consistent, indicating temperature of the tester electronics have reached steady state.

4.7.4 Pulse Insertion Unit Attachment

The ideal Pulse Insertion Unit (PIU) would be easy to install and would be produce exactly the same waveform from attachment to attachment. More precisely, there should be no effect of PIU installation on the collected data. To make a PIU performance judgment, a comparison is made between waveforms collected when the PIU is untouched between measurements and waveforms collected when the PIU has been reinstalled for each measurement. When first examining the PIU attachment repeatability after receiving the tester from the Pico-Second Pulse Laboratory, it was obvious that improvement was needed as the waveforms varied drastically for each PIU installation. Initially, the variation was thought to be due to the torque variations with the four PIU clamping screws. After testing various clamping torques, it became apparent that the order in which the four screws were tightened, in addition to the torque, was important. While this did improve the repeatability, the level of variation was still unacceptable.

Another attempt to deal with this problem was a redesign of the PIU. It was thought minute variations in the location of the DUT inside the PIU, due to tolerances in the alignment pin, could be the cause of the latent variations. The new PIU was designed so the coupling capacitance would not change if the DUT shifted or was installed in a slightly different location inside the PIU. While the new PIU did improve the repeatability, this design impacted the impedance matching between the TDR tester and the DUT, which introduced undesirable ringing in the resulting waveform. Therefore a different approach was required.

The solution to the PIU repeatability was to fabricate two thin shims to be installed between the PIU circuit plate and the PIU base plate. These shims are situated along side the cable so that the clamping screws can be tighten

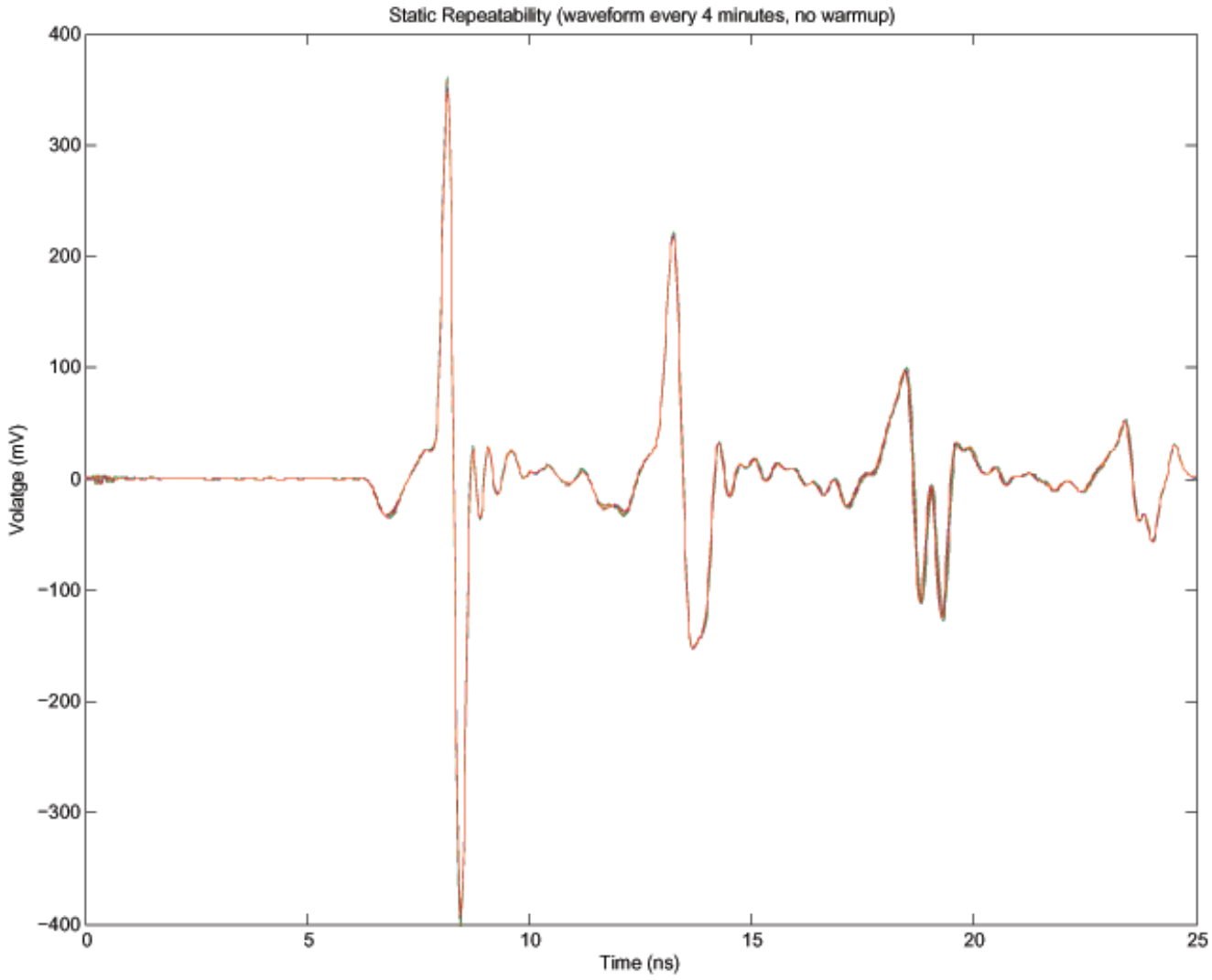


Figure 4.17: Static Repeatability: No warmup period before starting measurements. A total of ten waveforms were acquired, one every four minutes. Plots of the ten waveforms are overlaid in this figure.

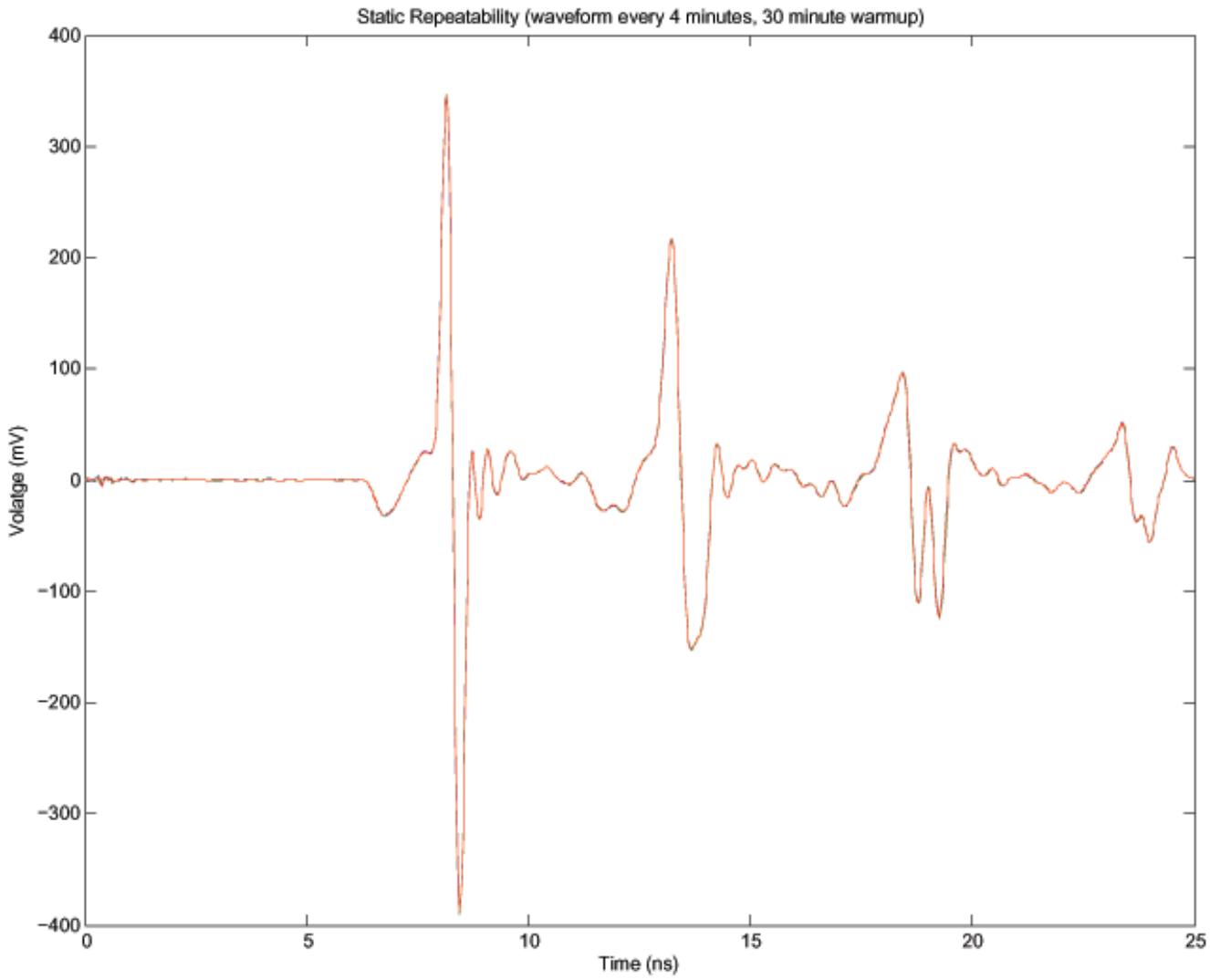


Figure 4.18: Static Repeatability: 30 minute warmup period before starting measurements. A total of ten waveforms were acquired, one every four minutes. Plots of the ten waveforms are overlaid in this figure.

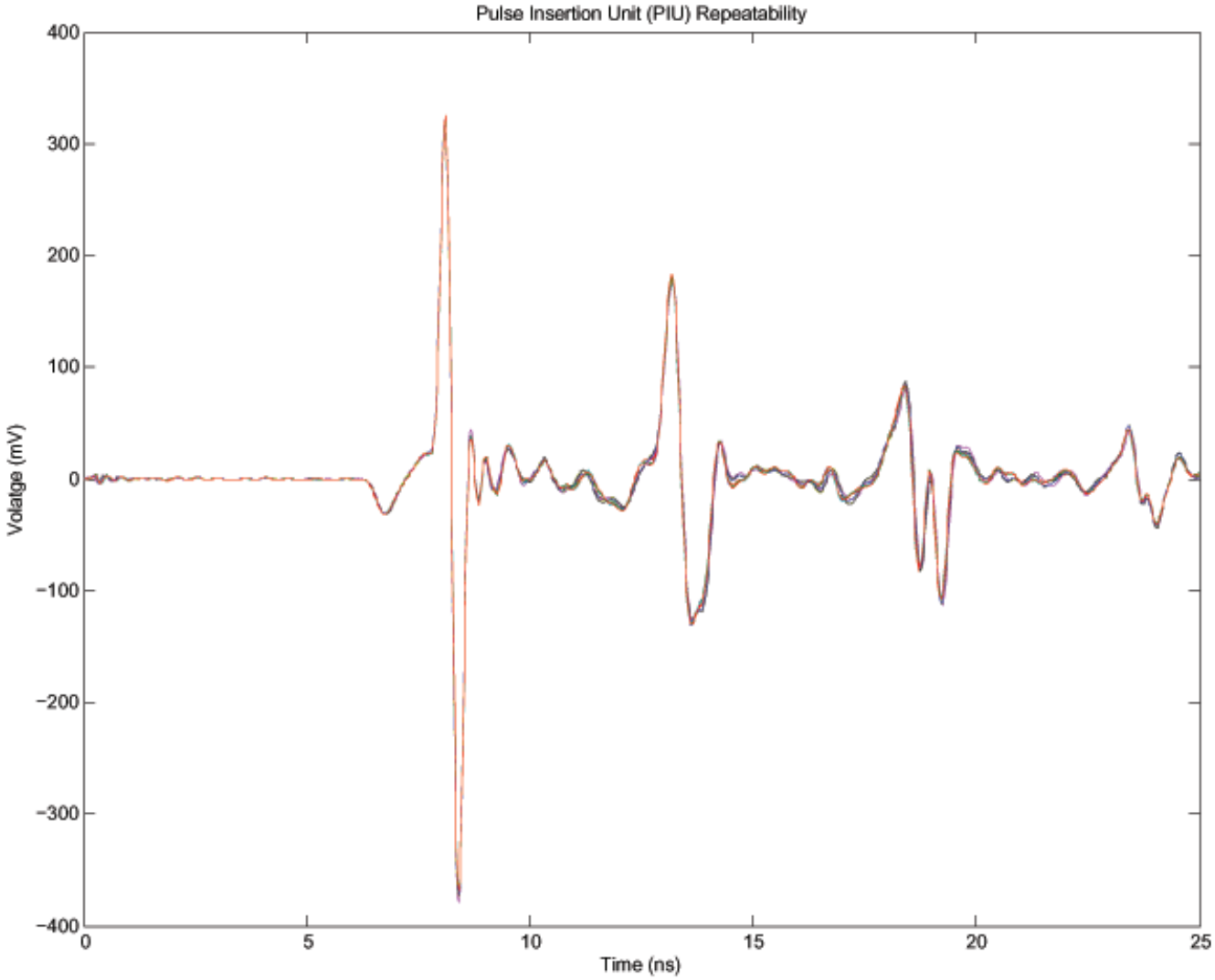


Figure 4.19: PIU Attachment Repeatability: The cable is attached to the PIU before each waveform is acquired. The cable is not removed from the test fixture between attachments. Ten waveforms were acquired in this manner and plotted above.

down completely onto the shims, compressing the cable exactly the same amount regardless of the clamping screw torque or tightening order. Various shim thickness were examined and 25 mils had the best results. Additionally, small solder mounds on the underside of the PIU circuit plate were noticed that needed to be made flush with the circuit board surface. These mounds introduced air gaps possibly affecting repeatability and introduced undesirable impressions on the cable. The final PIU attachment repeatability results can be seen in Figure (4.19) where ten waveforms were acquired, reinstalling the PIU between each measurement. The PIU design and assembly can be seen the TDR procedure AF-PT4194, which is including in the appendix.

4.7.5 2D Test Fixture Repeatability

The TDR waveform is very sensitive to the material surrounding the cable (especially metals that completely surround the cable) as well as the contour to which the cable conforms. Therefore it was necessary to design a test fixture that would simulate the physical environment of a cable in a weapon. In addition, the TDR test fixture should allow for easy installation and removal of cables in order to test many cables both before and after introducing damage.

These requirements resulted in the design of the 2D TDR test fixture. However, the process of installing a cable into the fixture introduces another set of processes that will affect the repeatability. Some of these variables include, cable location in the fixture, cable tension, and pressure applied to the various doors and covers. Each of these variables was examined individually and procedures and equipment were modified attempting to maximize the repeatability. For example, the fixture's aluminum cover initially was hinged in place and the weight of the cover was used to hold in it place. However, the repeatability study relating to this cover found this was introducing significant variation in the waveform. The fixture was then modified such that six screws, using a specified torque, hold the cover in place in order to make the installation and pressure more consistent.

The final repeatability variation introduced by the test fixture can be seen in Figure (4.20). To capture this variation, the cable was removed from the test fixture between measurements, but the PIU remained attached.

To capture the total system repeatability using the 2D test fixture, a single cable is tested, and between each measurement, the cable was removed from the test fixture and the PIU was detached. These results can be seen in Figure (4.21).

The variation that remains is thought to be a result to the manner in which the PIU is supported by the 2D test fixture. The two Teflon arms that hold the PIU allow some movement, so replacing them with a more rigid material is recommended. Additionally, making the PIU holding fixture from one piece, instead of two, would result in a more rigid and repeatable setup. These recommendations were used on the PIU holding fixture associated with the 3D Test Fixture.

4.7.6 Cable-to-Cable Repeatability in the 2D Test Fixture

It is somewhat difficult to capture the cable-to-cable repeatability, as any variation from the PIU attachment and fixture installation are superimposed on any data collected that requires a cable change. However, by comparing the variation from test series in which the cable is not changed Figure (4.21), to test series where the cable is changed Figure (4.22) we can get a qualitative estimate for the cable-to-cable variation. It should also be noted that the cable-to-cable variation dominates any variation from PIU attachment or test fixture installation.

However, unlike variation from the PIU attachment or test fixture installation, the cable-to-cable variation cannot be mitigated by improving the test process. Since this variation cannot be removed, this will limit the sensitivity of damage detection when comparing a cable of unknown condition to a typical good cable. If know good TDR waveforms exist for the specific cable that is now of unknown condition (a TDR birth certificate), the cable-to-cable variability becomes irrelevant.

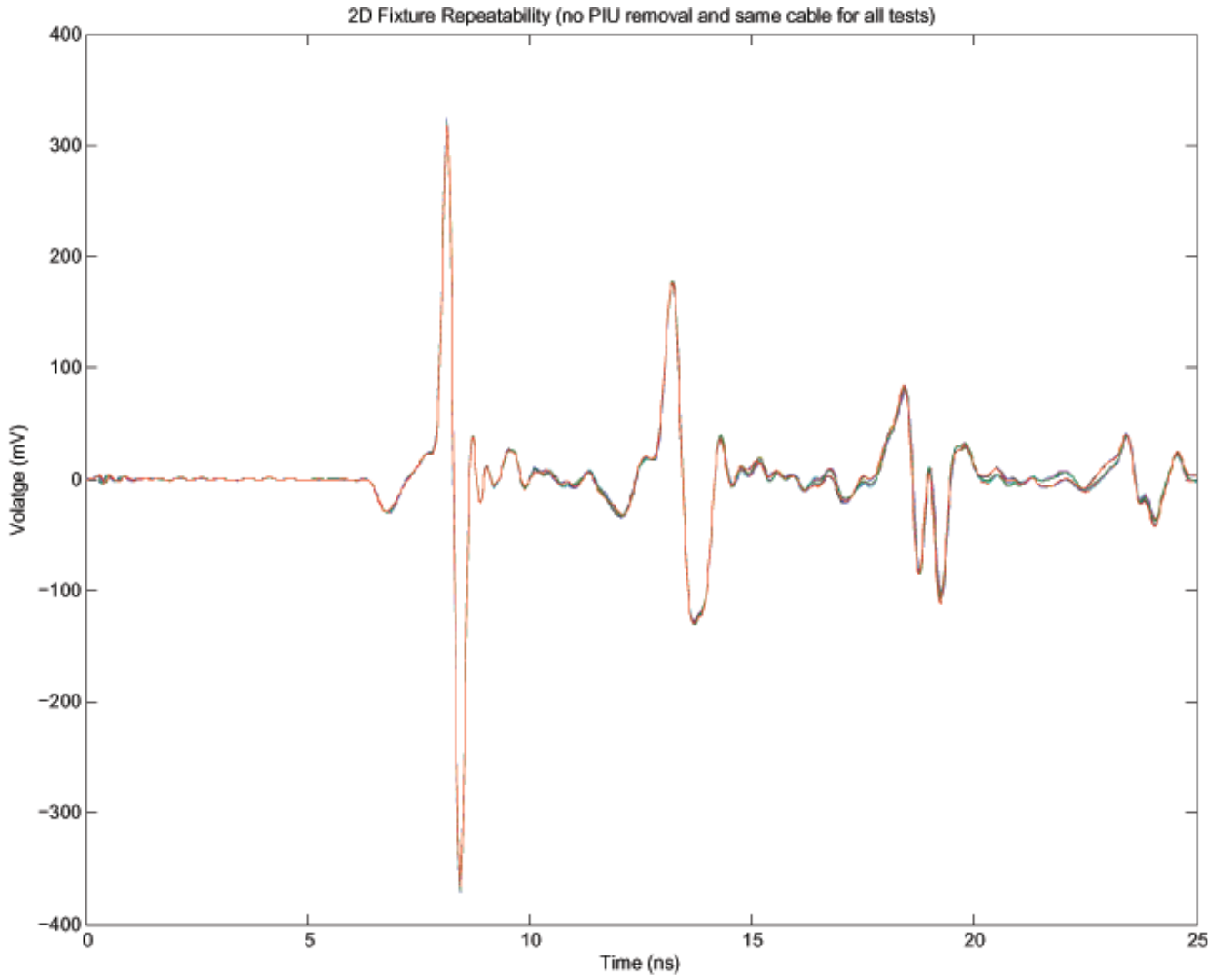


Figure 4.20: 2D Fixture Repeatability: The cable was removed from the test fixture between measurements, but the PIU remained attached. Ten waveforms were acquired in this manner and plotted above.

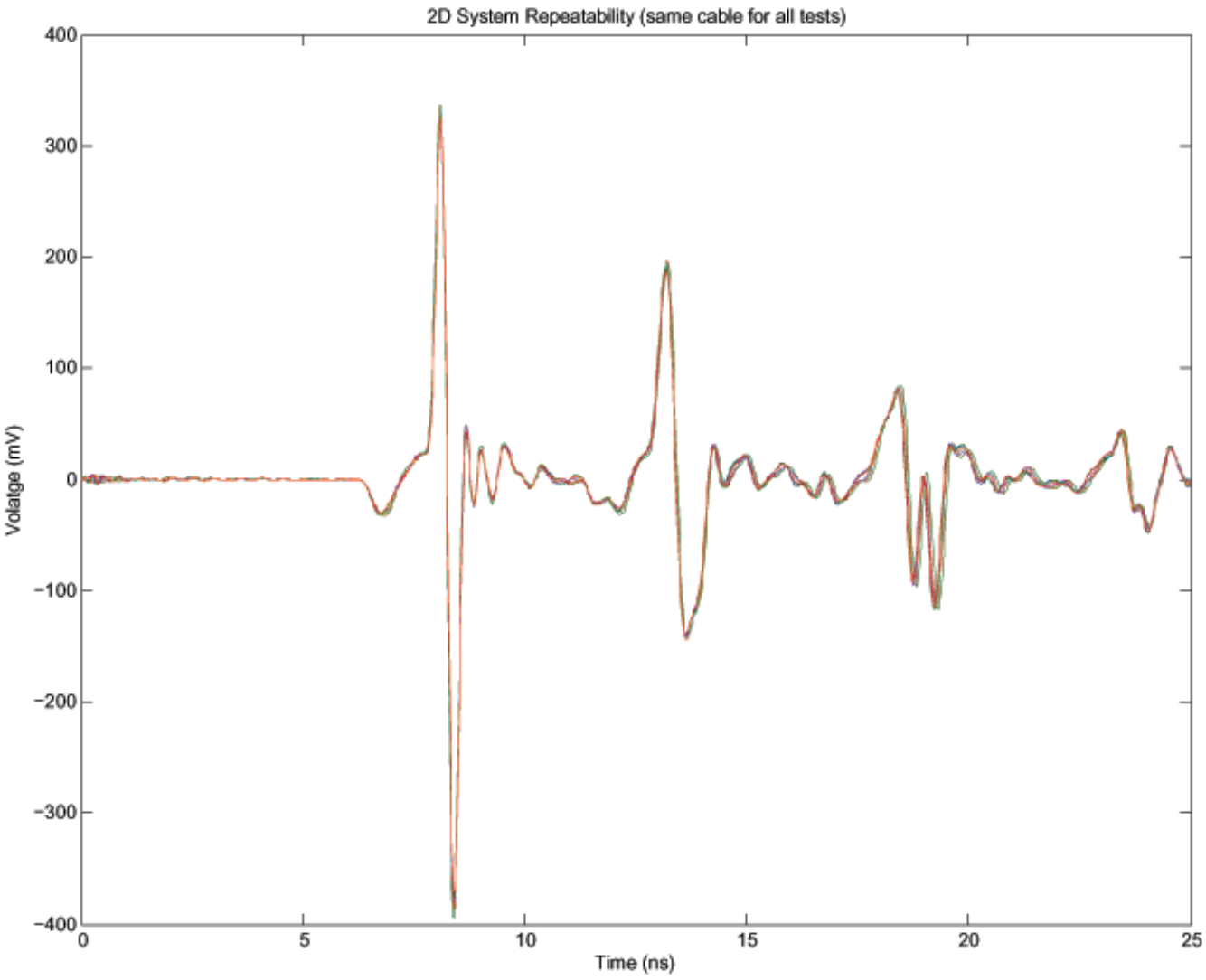


Figure 4.21: 2D System Repeatability: Between measurements, the cable was removed from the test fixture, and the PIU was detached. The same cable was used for all of the measurements. Ten waveforms were acquired in this manner and plotted above.

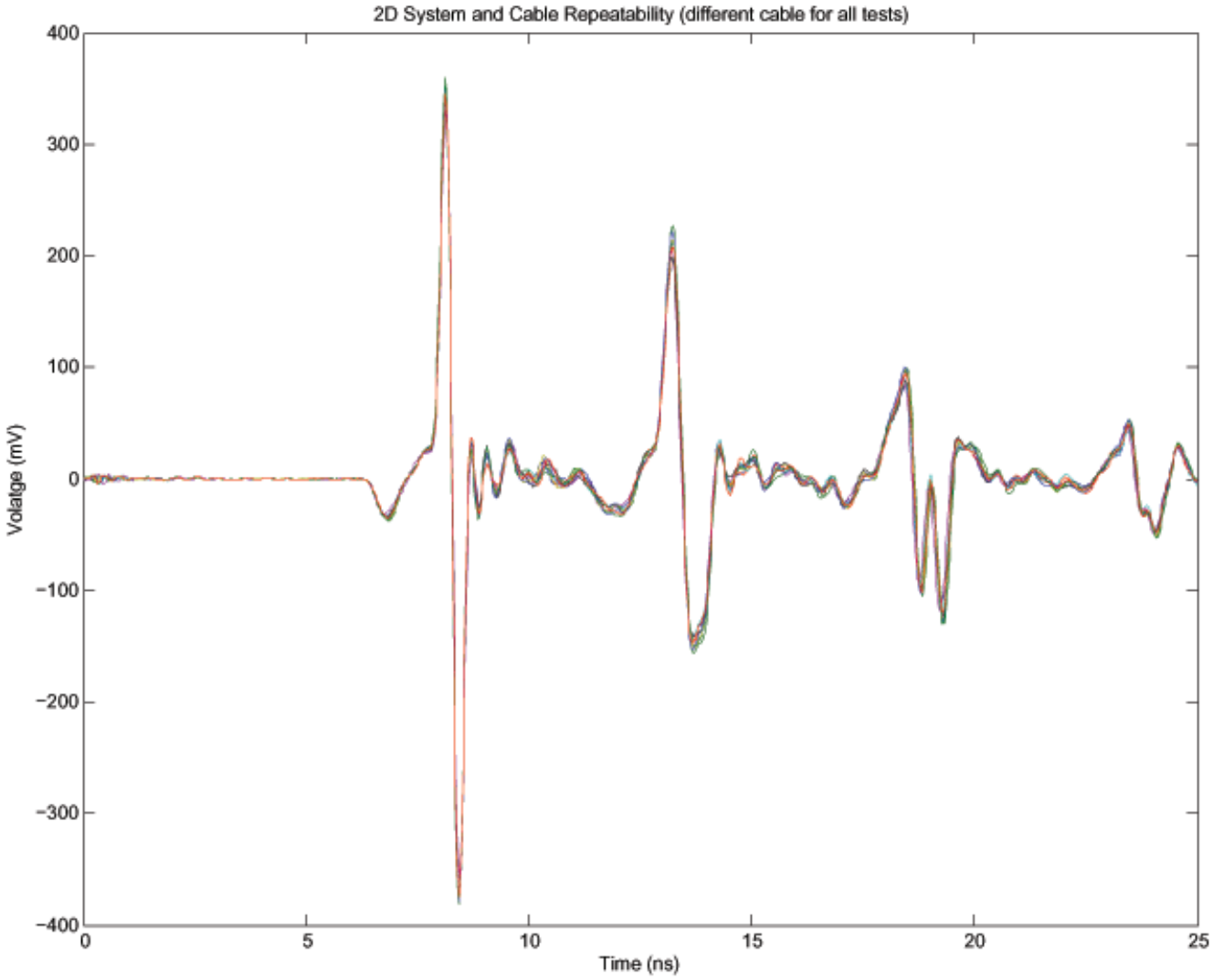


Figure 4.22: 2D System and Cable Repeatability: Between measurements, the cable was removed from the test fixture, and the PIU was detached. A different cable was used for each of the measurements. Ten waveforms were acquired in this manner and plotted above.

4.7.7 3D Test Fixture Repeatability

While the 2D test fixture will produce waveforms very similar to what is expected in a real weapon, there will be some difference as the 2D fixture design contains features to facilitate cable installation and removal. Additionally, surrogates were used for certain materials that could result minor differences. At this time is not possible to capture TDR data on a real weapon due the extensive safety analysis, documentation, and review that has not yet been finalized or completed. Therefore, in an attempt to capture a gold standard waveform to simulate a TDR signal from a good cable in real weapon, a high fidelity 3D test fixture was created and tested. Waveforms collected from the 3D test fixture can be seen in Figure (4.23) and have excellent repeatability. All tests with the 3D fixture utilize the same cable as the significant effort required to change out the cable makes doing so infeasible. The exceptional repeatability of these tests illustrates the benefits of acquiring TDR Birth Certificates, as doing so would make cable-to-cable and cable installation repeatability tests irrelevant.

4.7.8 Summary

This project has led to the following specific conclusions regarding repeatability:

(a) Once the PIU is attached to the cable, the TDR signal has very little variability from measurement to measurement.

(b) In the 2D fixture, from attachment to attachment for a single cable, the variability is significant, but reasonably small.

(c) From cable to cable, the variability is even greater, and presents a significant issue when damage detection must be made by comparison to typical good cable, as opposed to having a TDR birth certificate

An exhaustive effort and a large quantity of data were taken in order to work out and fine-tune the particular operations required to produce waveforms as repeatable as possible. There are literally dozens of variables associated with the PIU attachment and the DUT installation into the test fixture. Only after examining each of these variables independently and making tweaks to equipment and procedures, are highly repeatable data achievable.

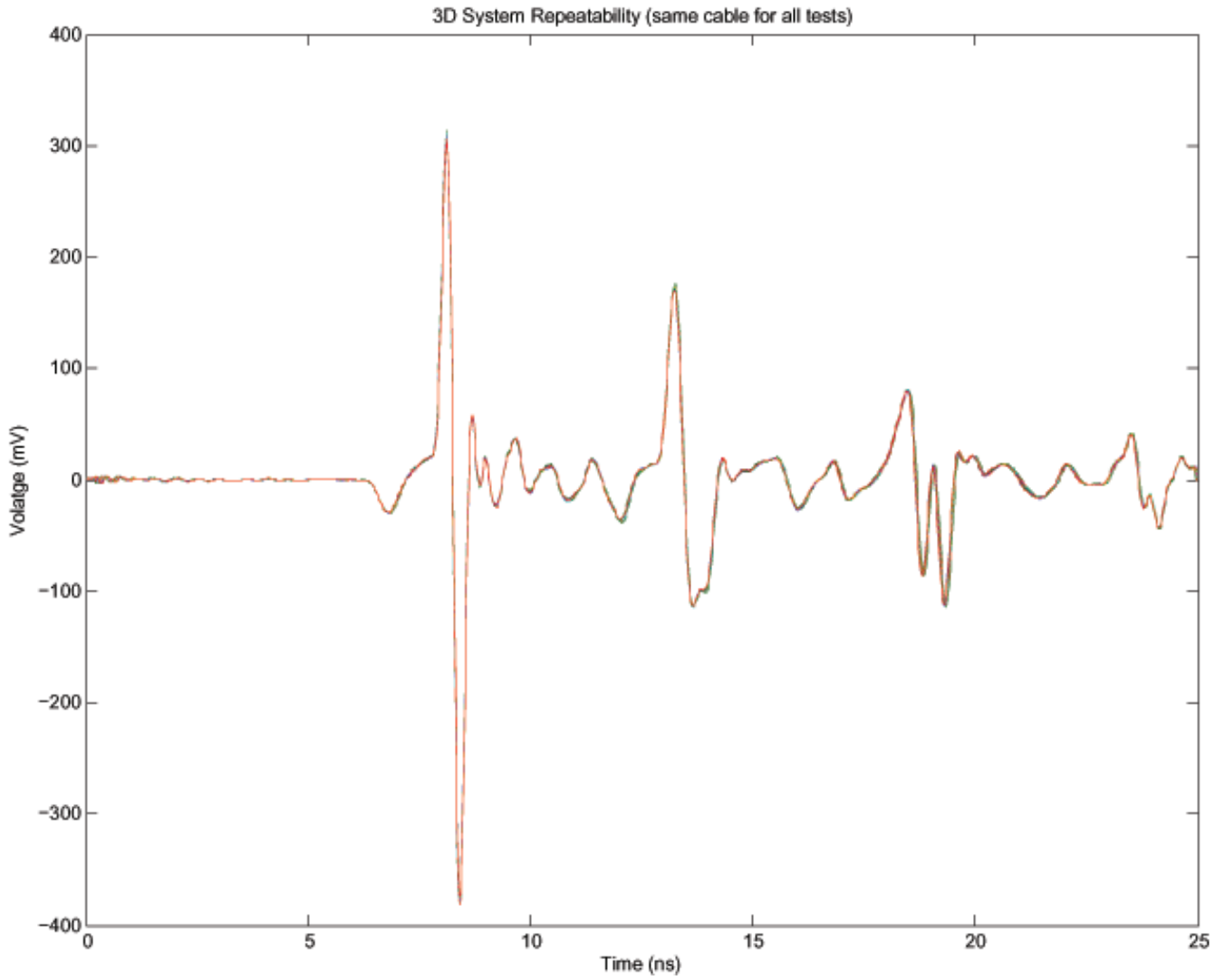


Figure 4.23: 3D System Repeatability: Between measurements, the PIU was detached, but the cable was not removed from the test fixture,. The same cable was used for all of the measurements. Ten waveforms were acquired in this manner and plotted above.

Chapter 5

Model-Based Damage Detection Algorithms

The model-based approach for detecting cable damage is depicted in block diagram form in Figures (5.1),(5.2) and (5.3). Two steps are involved. First, system identification algorithms are used to estimate a dynamic parametric model of the cable system, as in Figure (5.1). Here, signals for known undamaged cables are used to estimate the model. Second, the model is used in comparing the signal for a cable under test with the signal from the model as in Figure (5.2). Both steps are summarized together in Figure (5.3).

Step 1, Cable Modeling: Given input-output TDR signals $s(n)$ and $x_U(n)$ for a cable known to be free of damage, system identification algorithms are used to compute a dynamic prediction-error cable model that has output $\widehat{x}_U(n)$. The model is declared valid when the innovations $e_U(n) = x_U(n) - \widehat{x}_U(n)$ satisfy a statistical zero-mean whiteness test. This validated model output is then used as a known reference to which other cables can be compared.

Step 2, Cable Testing: The TDR output signal $x_D(n)$ from a cable under test is compared with the model output $\widehat{x}_U(n)$ by computing the innovations $e_D(n) = x_D(n) - \widehat{x}_U(n)$. The innovations are tested using a short-term whiteness test statistic, which employs a statistical confidence interval. If the cable passes the test, this implies that the model is valid and the cable is declared undamaged. If the cable fails the test, this indicates a model mismatch, which means that the cable's dielectric properties have changed; and this implies that the cable is damaged. The test threshold is adjusted to maximize probability of detection and minimize probability of false alarm according to an empirically determined receiver operating characteristic (ROC) curve. An associated confidence interval on the probability of correct classification is also provided.

5.1 Parametric Model for the Cable Dynamics

We choose to use a parametric prediction error model for the cable dynamics. This is because such models have the very desirable property that the prediction errors (innovations) must be white when the model fits the measurements. We show in a later section how the innovations can be used to create a very effective model mismatch detector that infers the functionality of the cable.

Let t_n denote time at discrete time index n. Now let us define an autoregressive model with exogenous inputs (ARX), which takes the general form $ARX(N_a, N_b)$ as follows ([43, 44, 45, 52, 53, 54]. The model is in the form of a linear difference equation.

$$y(t_n) + a_1 y(t_n - 1) + \ldots + a_{N_a} y(t_n - N_a) = b_1 u(t_n - 1) + \ldots + b_{N_b} u(t_n - N_b) + e(t) \tag{5.1}$$

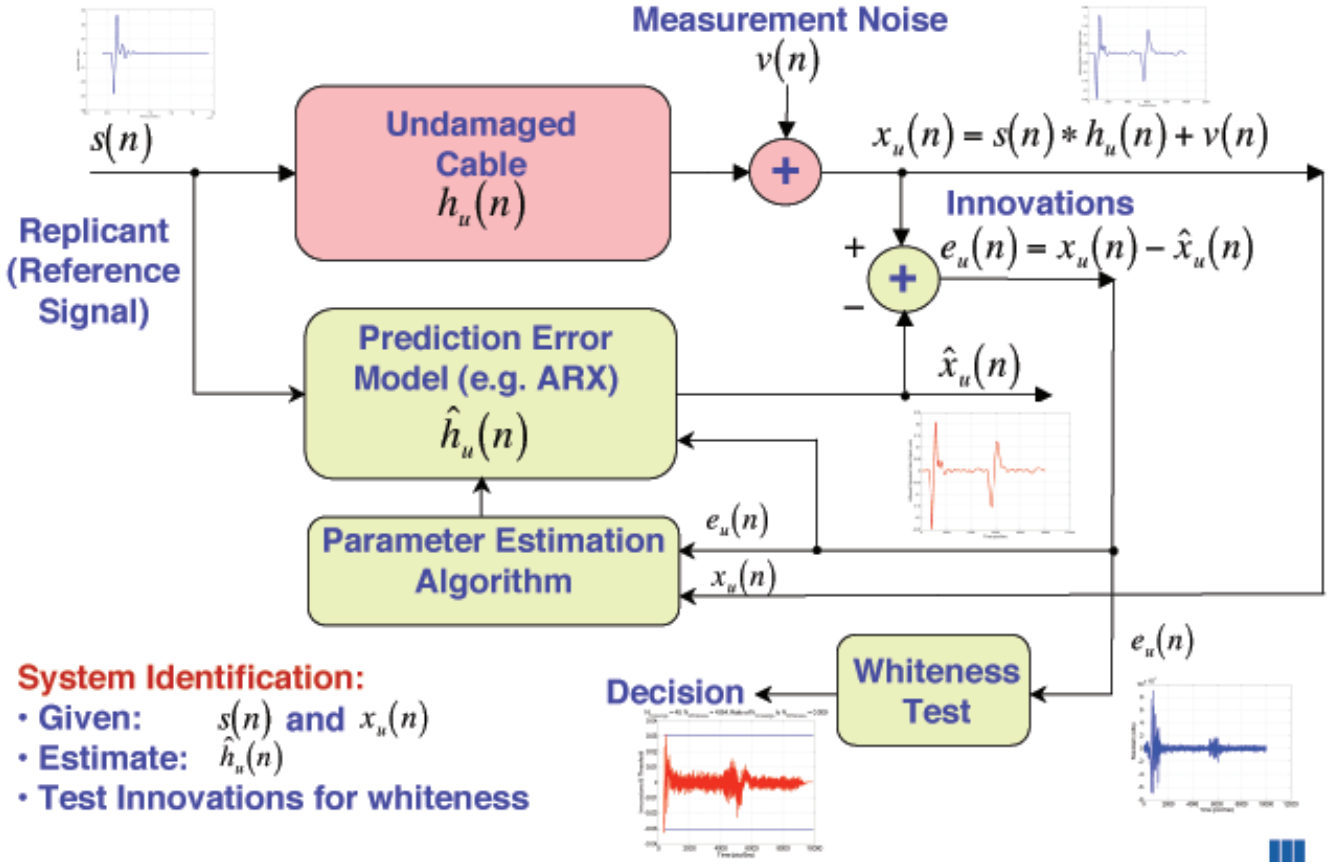


Figure 5.1: Step 1: The first step in the damage detection algorithms is system Identification to estimate a parametric prediction error model for the cable under test. The measured input $s(n)$ and output $x_U(n)$ are used to estimate the parameters of an ARX model and produce the estimated model output $\widehat{x}_U(n)$. The innovations are then the difference signal $e(n) = x_U(n) - \widehat{x}_U(n)$. If the innovations pass the zero-mean whiteness test, then the model is declared to be valid.

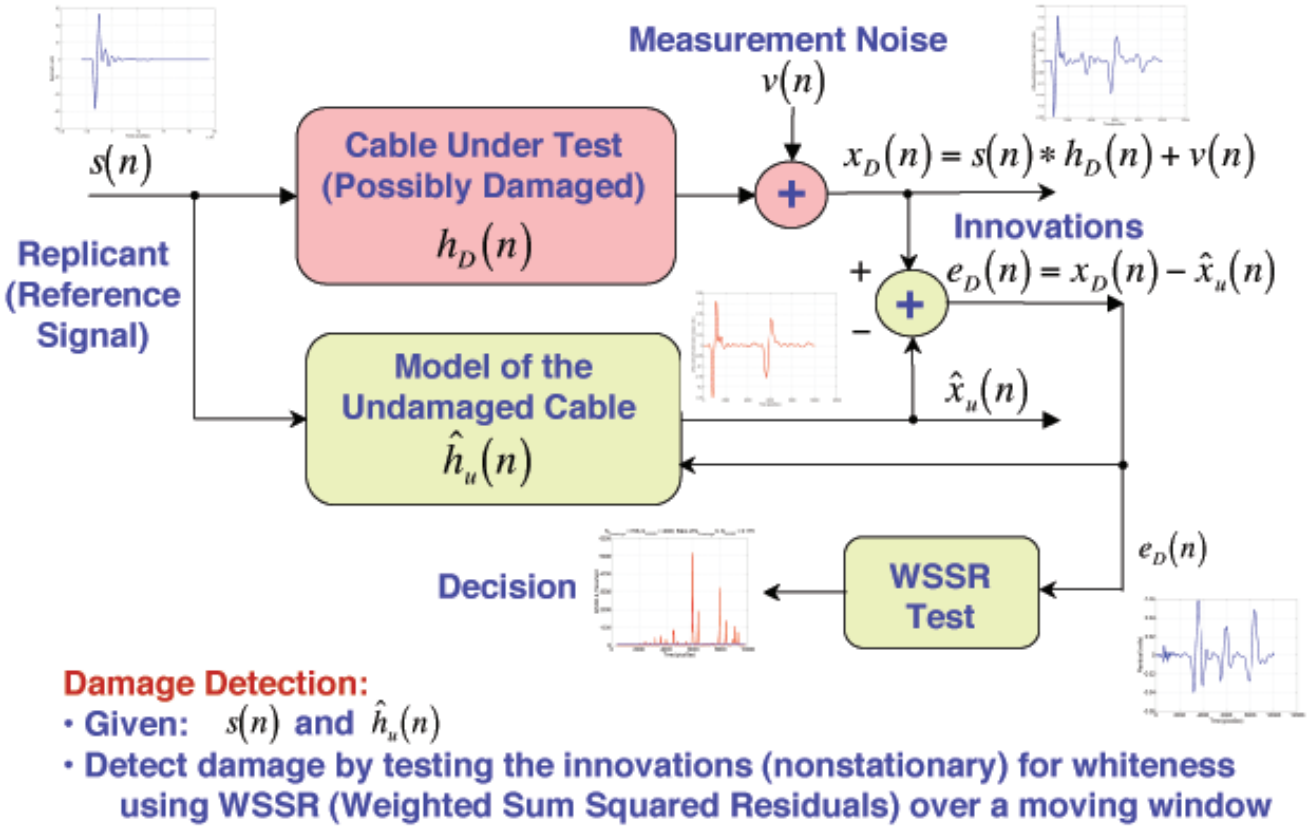


Figure 5.2: Step 2: The second step in the damage detection algorithms is to compare a signal $x_D(n)$ from a cable that is possibly damaged (device under test) against the estimated model output signal $\widehat{x}_U(n)$. This is done by forming the test innovations $e_D(n) = x_D(n) - \widehat{x}_U(n)$ and testing them with the WSSR test statistic, which is a short-term whiteness statistic. If the test innovations pass the WSSR test, then no model mismatch is detected, and the cable is declared undamaged. If the innovations fail the WSSR test, then a model mismatch is detected and the cable is declared damaged.

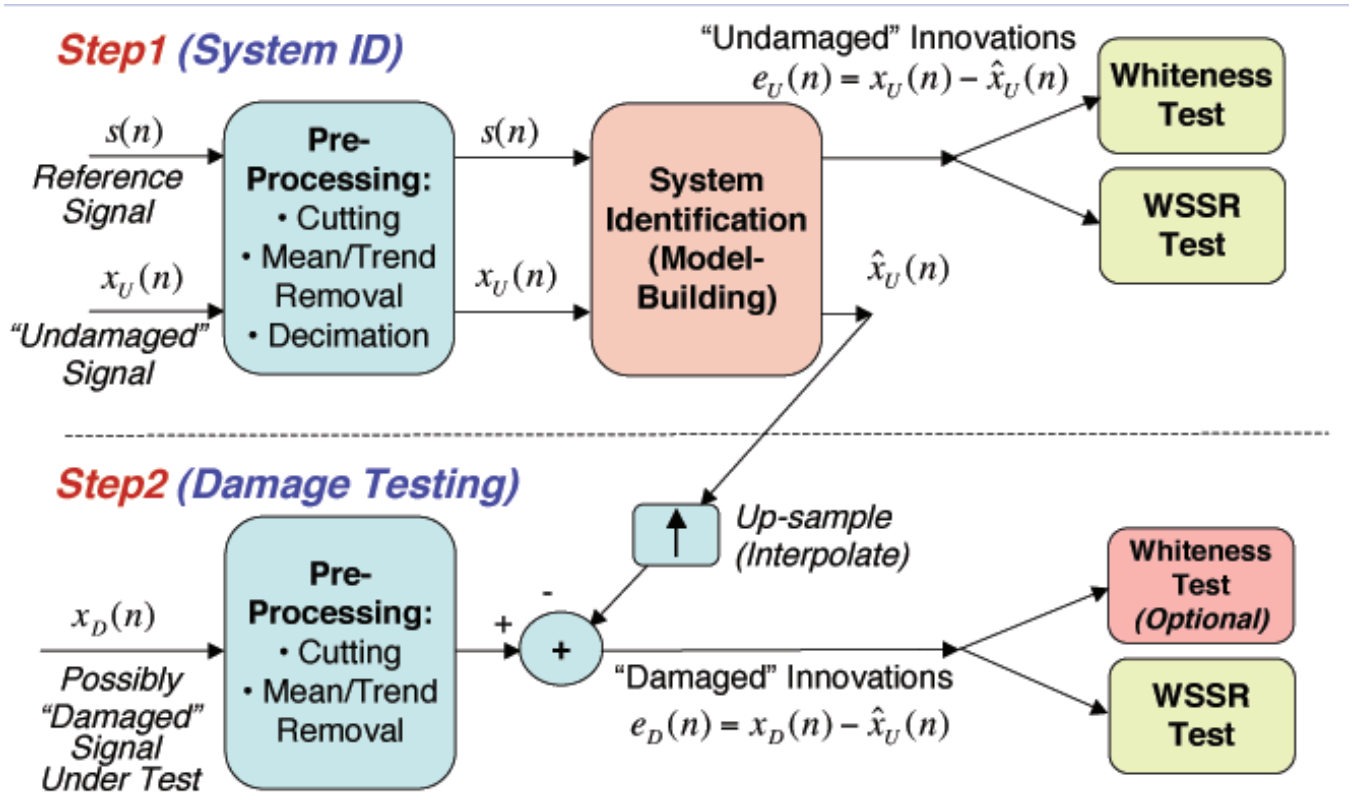


Figure 5.3: Step 1: The Steps 1 and 2 described in the previous two figures are summarized in this figure. The first step is to identify a cable model using a reference input and output from a cable that is known to be undamaged. The second step is to test a new signal from a cable that is possibly damaged using the WSSR test on the test innovations $e_D(n) = x_D(n) - \widehat{x}_U(n)$.

where $y(t_n)$ is the measurement signal and $e(t_n)$ is zero-mean, white Gaussian noise, $e(t) \sim N(0, \sigma^2{}_e)$. The variance of the Gaussian distribution is denoted by σ^2. Let us define the backward shift (or delay) operator q^{-i}, where i denotes an integer index:

$$q^{-i}y(t_n) = y(t_{n-i}) \tag{5.2}$$

Now we can define instantaneous polynomials at time t_n at follows:

$$A(q^{-1}, t_n) = 1 + a_1(t_n)q^{-1} + ... + a_{N_a}(t_n)q^{-N_a} \tag{5.3}$$

$$B(q^{-1}, t_n) = b_1(t_n)q^{-1} + ... + b_{N_b}(t_n)q^{-N_b} \tag{5.4}$$

Using this notation, we can write the parametric dynamic model in difference equation form as follows:

$$A(q^{-1}, t_n)y(t_n) = B(q^{-1}, t_n)u(t_n) + e(t_n) \qquad \text{ARX Model} \tag{5.5}$$

It is convenient and practical to adjust the model to account for a general delay that commonly exists between the input and output in real systems. For this reason, we include a delay of N_k samples in the input measurement as follows:

$$A(q^{-1}, t_n)y(t_n) = B(q^{-1}, t_n)u(t_n - nN_k) + e(t_n) \qquad \text{ARX Model} \tag{5.6}$$

Rearranging terms, we obtain another form of the model:

$$y(t) = \frac{B(q^{-1}, t_n)}{A(q^{-1}, t_n)}u(t_n - nN_k) + \frac{1}{A(q^{-1}, t_n)}e(t) \qquad \text{ARX Model} \tag{5.7}$$

This model is denoted by $ARX(N_a, N_b, N_k)$ and is depicted in block diagram form in Figure (5.4). We can cast this model into vector form by defining the following d x 1 measurement and parameter vectors, where $d = N_a + N_b + N_k$.

$$\underline{\phi}(t_n) = [-y(t_n - 1) \cdots - y(t_n - N_a)|u(t_n - N_k) \ldots u(t_n - N_k - N_b + 1)]^T \quad \text{measurement vector} \tag{5.8}$$

$$\underline{\theta} = [a_1 \ldots a_{N_a}|b_1 \ldots b_{N_b}]^T \qquad \text{parameter vector} \tag{5.9}$$

For the estimation problem, it can be shown [52] that the conditional expectation of the measurement $y(t)$ given the parameter vector $\underline{\theta}$ is given by the following:

$$\widehat{y}(t_n|\underline{\theta}) = B(q^{-1}, t_n)u(t_n) + \left[1 - A(q^{-1}, t_n)\right] y(t_n) \qquad \text{Measurement Estimate} \tag{5.10}$$

Substituting the definitions above, we can put this estimate in vector form as follows:

$$\underline{\widehat{y}}(t|\underline{\theta}) = \underline{\phi}^T(t_n)\underline{\theta} \qquad \text{Measurement Estimate} \tag{5.11}$$

Now we can define the prediction error as follows:

$$\varepsilon(t_n, \underline{\theta}) = y(t_n) - \underline{\widehat{y}}(t_n|\underline{\theta}) \tag{5.12}$$

$$= y(t_n) - \underline{\phi}^T(t_n)\underline{\theta} \qquad \text{Prediction Error} \tag{5.13}$$

The least squares criterion becomes:

$$V_N(\underline{\theta}) = \frac{1}{N}\sum_{t_n=1}^{N}\frac{1}{2}\left[y(t_n) - \underline{\phi}^T(t_n)\underline{\theta}\right]^2 \qquad \text{Least Squares Criterion} \tag{5.14}$$

This criterion has the desirable property that is quadratic in the parameter vector $\underline{\theta}$. This means that it can be minimized analytically. The Least Squares Estimate is defined as:

$$\underline{\widehat{\theta}}_N^{LS} = \arg\min\left[V_N(\underline{\theta})\right] \qquad \text{Least Squares Estimate} \tag{5.15}$$

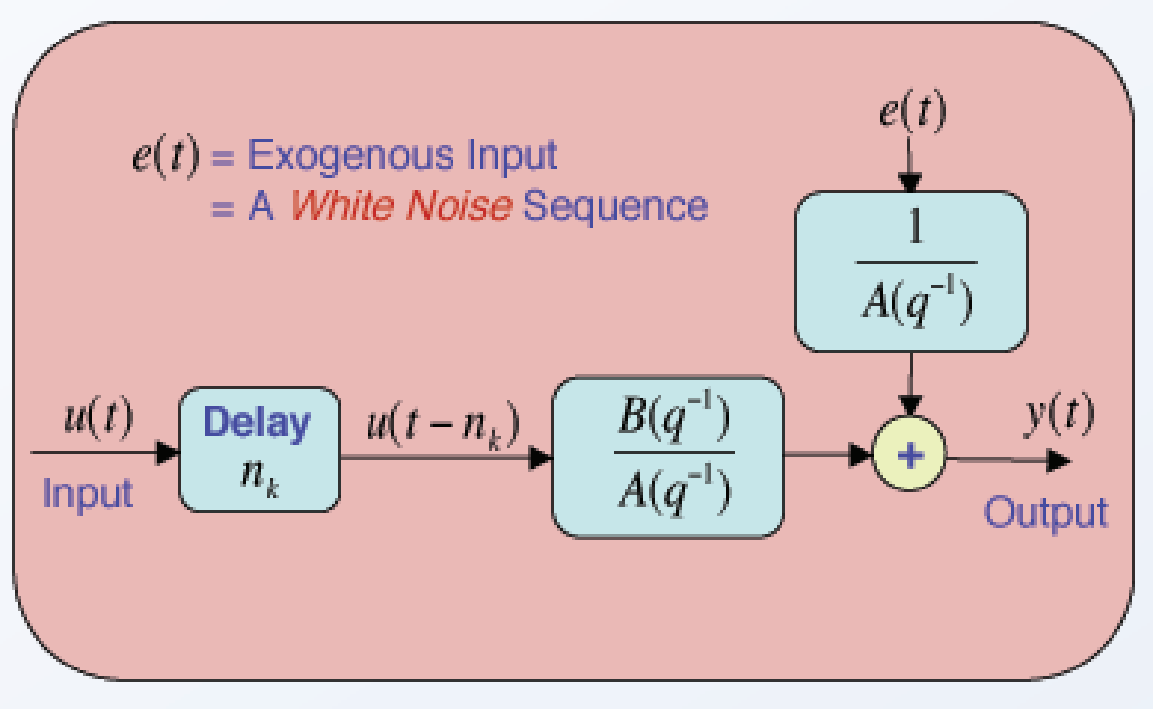


Figure 5.4: Step 1: The ARX model $ARX(N_a, N_b, N_k)$ of Equation (5.7) is depicted here in block diagram form.

Doing the minimization analytically, we obtain the set of Normal Equations which are solved to obtain the Least Squares estimate of the parameters:

$$\left[\frac{1}{N}\sum_{t_n=1}^{N}\underline{\phi}(t)\underline{\phi}^T(t)\right]\widehat{\theta}_N = \left[\frac{1}{N}\sum_{t_n=1}^{N}\underline{\phi}^T(t)y(t_n)\right] \qquad \text{Normal Equations} \qquad (5.16)$$

Rearranging the Normal Equations gives the Least Squares estimate of the parameters:

$$\widehat{\underline{\theta}}_N^{LS} = \left[\frac{1}{N}\sum_{t_n=1}^{N}\underline{\phi}(t)\underline{\phi}^T(t)\right]^{-1}\left[\frac{1}{N}\sum_{t_n=1}^{N}\underline{\phi}^T(t)y(t_n)\right] \qquad \text{LS Estimate} \qquad (5.17)$$

Let us now write the least squares problem in vector form, because this is how it is actually solved in practice. Define the N x 1 column vector of system output measurements over a block of time that is N samples long.

$$\underline{Y}_N = [y(1)\ldots y(N)]^T \qquad \text{N x 1 Output vector} \qquad (5.18)$$

Next, define an N x d matrix in which each row is a 1 x d measurement vector at a given time sample over the block of time N samples long.

$$\Phi_N = \begin{bmatrix} \underline{\phi}^T(1) \\ \underline{\phi}^T(2) \\ \vdots \\ \underline{\phi}^T(N) \end{bmatrix} \qquad \text{N x d Matrix of Measurements} \qquad (5.19)$$

Using this notation, we can write the measurement equation in matrix-vector form as follows:

$$\underline{Y}_N = \Phi_N\underline{\theta} \qquad \text{Measurement Equation} \qquad (5.20)$$

where $\underline{Y}_N$ is (N x 1), Φ_N is (N x d), and $\underline{\theta}$ is (d x 1). Corresponding to Equation (5.11), the conditional estimate of $\underline{Y}_N$ based upon the parameters is given by:

$$\underline{\widehat{Y}}_N(\underline{\theta}) = \Phi_N \underline{\theta} \qquad \text{Conditional Output Estimate} \tag{5.21}$$

The vector prediction error then, is:

$$E_N = \underline{Y}_N - \underline{\widehat{Y}}_N(\underline{\theta}) \qquad \text{Vector Prediction Error} \tag{5.22}$$

The Least Squares Criterion then becomes:

$$V_N(\underline{\theta}) = \frac{1}{N} \left| \underline{Y}_N - \underline{\widehat{Y}}_N(\underline{\theta}) \right|^2 \qquad \text{Least Squares Criterion} \tag{5.23}$$

The normal equations are:

$$\left[\Phi_N^T \Phi_N\right] \widehat{\theta}_N = \Phi_N^T \underline{Y}_N \qquad \text{Normal Equations} \tag{5.24}$$

Finally, the normal equations can now be rearranged using the pseudo-inverse to form the Least Squares estimate of the parameter vector:

$$\underline{\widehat{\theta}}_N^{LS} = \left[\Phi_N^T \Phi_N\right]^{-1} \Phi_N \underline{Y}_N \qquad \text{Least Squares Estimate} \tag{5.25}$$

For the work in this report, the ARX model parameters are estimated in MATLAB using the least squares algorithm described above. It solves an over-determined set of linear equations using QR factorization. The regression matrix is formed so that only measured quantities are used (no fill-out zeros) ([68, 69]).

5.2 Whiteness Test on the Innovations

We saw in the discussion on system identification that the innovations sequence is tested for whiteness to determine whether or not the estimated model fit is acceptable. This section describes how the whiteness test is developed and performed for a single innovations signal.

We define two hypotheses, H_0 and H_1 as follows:

H_0 : The innovations are white (the null hypothesis is true), so the model matches the data.
H_1 : The innovations are not white, so a model mismatch exists, so the model needs adjustment. Assuming that the innovations are ergodic, we use the normalized sample covariance as a test statistic:

$$\widehat{\rho}_e(k) = \frac{\widehat{R}_e(k)}{\widehat{R}_e(0)} \qquad \text{Test Statistic} \tag{5.26}$$

Where k is the correlation lag index, M is the number of samples in the innovtions sequence, the asterisk denotes complex conjugation and the sample covariance is:

$$\widehat{R}_e(k) = \begin{cases} \frac{1}{M} \sum_{n=k_1}^{M-|k|-1} \left[e(n) - \widehat{m}_e\right] \left[e(n+k) - \widehat{m}_e\right] \\ {R^*}_e(-k) \end{cases} \tag{5.27}$$

and the sample mean is:

$$\widehat{m}_e = \frac{1}{M} \sum_{n=1}^{M} e(n) \tag{5.28}$$

Note: We sum from $n = k_1$ on to avoid the first sample (at zero lag) which would equal one when we apply the hypothesis test described next. It can be shown that the test statistic is Chi Squared distributed with M degrees of freedom $\gamma(n) \sim \chi^2(W)$ [43]. Asymptotically, for large M, it can be shown that the test statistic is Gaussian distributed [23] as follows. In practice, M is generally considered to be large if $M > 30$.

$$\widehat{\rho}_e(k) \sim N\left[0, \frac{1}{M}\right] \qquad \text{for } M > 30 \tag{5.29}$$

Therefore, it can be shown [66] that we can write a 95% confidence interval estimate of the test statistic as:

$$I_{\rho_e} = \widehat{\rho}_e(k) + \frac{1.96}{\sqrt{M}} \qquad \text{for } M > 30 \tag{5.30}$$

Note that the sample size M must be large ($M > 30$) for the confidence interval to be accurate [66]. Another way to express the confidence interval is:

$$P\left\{\widehat{\rho}_e(k) - \frac{1.96}{\sqrt{M}} < \rho_e(k) < \widehat{\rho}_e(k) + \frac{1.96}{\sqrt{M}}\right\} = 1 - \alpha = .05 \tag{5.31}$$

where α is the significance level of the test [66]. We read this statement as follows, "with confidence 95%, the test statistic lies between the lower bound and the upper bound." So, under the null hypothesis, the test statistic $\rho_e(k)$ must lie within the confidence interval I_{ρ_e} 95% of the time to declare that the innovations are white.

In practice, we implement the whiteness test as follows: Count the number of samples N_e of $\widehat{\rho}_e(k)$ that exceed the thresholds, and compute the percentage of samples that exceed the tresholds, $F_e = [N_e/(M - k_1)]$. If $F_e \leq \alpha$, declare H_0 is true (innovations are white). If $F_e > \alpha$, declare H_1 is true (innovations are not white).

5.3 WSSR Test on the Innovations

When the signals are nonstationary, the Weighted Sum Squared Residuals (WSSR) criterion is a more reliable statistic to use for testing the whiteness of the prediction error sequence (innovations). For the cable damage detection problem, we need to assume that the innovations can be non-stationary, because any damage will cause a disturbance in the innovations. Thus, WSSR is a reasonable choice to use as a decision statistic for detecting damage. WSSR has the desirable property that is is good for detecting "jumps "or discontinuities. For model-based damage detection, we require that the WSSR lies beneath a calculated threshold to deem the innovations zero-mean and white. WSSR is calculated over a finite sliding window of W samples as depicted in Figure (5.5). Under the zero mean assumption, the WSSR statistic is equivalent to testing that the prediction error sequence is white. We define the WSSR test statistic (a scalar) as follows:

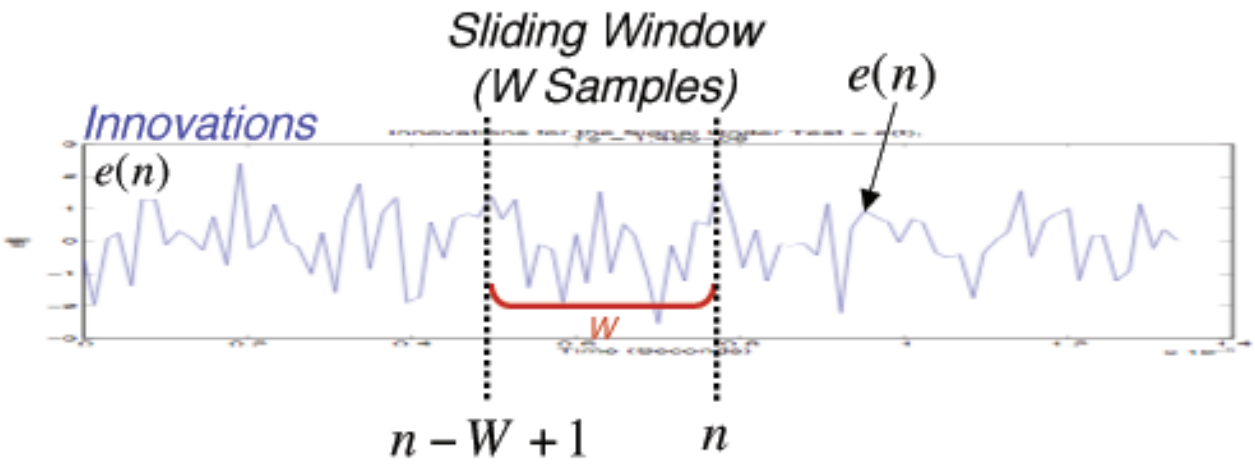
$$\gamma(n) = \sum_{j=n-W+1}^{n} \frac{e^2(j)}{V(j)} \qquad \text{for } n \geq W \tag{5.32}$$

Note: we evaluate the WSSR over a sliding window of length W samples. We calculate $\gamma(n)$ only for $n \geqslant W$. This means that we must "waste "W samples at the beginning. This is a good reason to ensure that the measured signals x_u and x_d contain plenty of pre-event background. The autocovariance is defined as follows:

$$V(n) = \frac{1}{W} \sum_{j=n-W+1}^{n} \left[e^2(j) - \bar{e}(j)\right]^2 \qquad \text{for } n \geq W \tag{5.33}$$

and the sample mean over the window is:

$$\bar{e}(n) = \frac{1}{W} \sum_{j=n-W+1}^{n} e(j) \qquad \text{for } n \geq W \tag{5.34}$$



$$\gamma(n) = \sum_{j=n-W+1}^{n} \frac{e^2(j)}{V(j)}, \quad \text{for } n \geq W$$

Figure 5.5: WSSR (Weighted Sum Squared Residuals) is calculated over a sliding window W samples wide.

By defining a threshold τ, the WSSR test becomes:

$$\text{if } \gamma(n) \underset{H_0}{\overset{H_1}{\gtrless}} \tau \tag{5.35}$$

Under the null hypothesis, the WSSR is distributed $\gamma(n) \sim \chi^2(W)$. However, for large W ($W > 30$), the WSSR is approximately normally distributed:

$$\gamma(n) \sim N(W, 2W) \qquad \text{for } W > 30 \tag{5.36}$$

At the significance level α, it can be shown that the probability of rejecting the null hypothesis is [44, 66]:

$$P\left\{\left|\frac{\gamma(n) - W}{\sqrt{2W}}\right| > \left|\frac{\tau - W}{\sqrt{2W}}\right|\right\} = \alpha \tag{5.37}$$

For a significance level of $\alpha = .05$, the test threshold is:

$$\tau = W + 1.96\sqrt{2W} \tag{5.38}$$

and another way to write the WSSR test on the innovations is:

$$H_0 : P\{\gamma(n) < \tau\} = 1 - \alpha \tag{5.39}$$
$$H_1 : P\{\gamma(n) \geqslant \tau\} = \alpha \tag{5.40}$$

In practice, we implement the WSSR test by calculating the WSSR test statistic $\gamma(n)$ for a chosen window length W. We then count the fraction F_E of WSSR samples that exceed the threshold. If $F_E > \alpha$, we declare that there exists a model mismatch. If $F_E \leqslant \alpha$, we declare that there exists no model mismatch.

5.4 WSSR Test on Signal Mean Residuals

We implemented an alternative ad hoc WSSR test to provide another option for cable damage detection. We considered the option of using the WSSR with a signal mean residual signal instead of the model-based innovations. The signal mean residuals are defined as follows:

$$r(n) = x_D(n) - x_{Uavg}(n) \qquad \text{Signal Mean Residuals} \qquad (5.41)$$

where $x_{Uavg}(n)$ represents the sample ensemble mean of a set of TDR signals measured for a set of undamaged cables. This residual represents an average reference signal for undamaged cables. In fact, the signal $x_{Uavg}(n)$ is the same signal we used to identify a cable dynamic model for the model-based scheme (see the next chapter on signal pre-processing).

The advantage of using $x_{Uavg}(n)$ is that it is a good representation of the undamaged class of signals. The disadvantage of using $x_{Uavg}(n)$ is that it does not have the nice theoretical properties that the innovations have. If the model is valid, the innovations must be white. However, there is no such requirement for the residuals. The residuals may or may not be correlated - but in general, they will NOT be correlated.

Nonetheless, we decided that we would like to apply the WSSR test to the residuals to test performance. Unfortunately, project resources limited the time we could spend on this path. We ran one simple test case, and found that the WSSR is able to detect obvious cable damage when applied to the signal average residual. However, more work would need to be done to assess performance under more challenging scenarios.

Chapter 6

Signal Preprocessing Before System Identification and Detection

6.1 Pre-Processing for the System Identification Step

Signal preprocessing is extremely important to the practical success of the system identification algorithms. The preprocessing steps ensure that the the signals satisfy the algorithm assumptions and give the algorithm the best chance to overcome the fact that the basic mathematical problem is ill-posed [15]. In addition, the preprocessing steps allow us to put the data in a form that is most easily interpreted.

Several preprocessing steps are executed for the System Identification Step. These steps are briefly introduced here, then described in greater detail with examples in the following subsections.

(1) **Auto-Alignment** of the raw signals to be used for system identification. See the earlier discussion on auto-alignment.

(2) **Cutting** the desired signal out of the raw signal. Here, we wish to use only the first two cable reflections, because the rest are not useful for damage detection.

(3) **Ensemble averaging** of J signals each to form the $s(n)$ and $x_U(n)$ signals to use for system identification. For our problem, we most often chose $J = 3$ because of practical constraints from the program. We have many undamaged cables, but we have only one long cable. Also, at the time of some of our studies, we had only a few signals for our use.

(4) **Creating a Reference Signal** $s(n)$ by cutting a longer signal, then pre-and-post-padding it with zeros to ensure that it has the same length as $x_U(n)$. See the section below.

(5) **Temporal Shifting** of the $s(n)$ and/or the $x_U(n)$ signal to ensure causality for the system to be identified. The cable is a causal device, so the signals used to build the model must be causal. For the cable problem, this means that the input signal $s(n)$ must start (become nonzero) before the cable output signal $x_U(n)$ starts. This is obvious, because the output cannot occur before the input occurs (causality).

(6) **Removing the Mean and any Linear Trend** from the signals to ensure that we satisfy the linearity assumption. This must be done because the cable is assumed to be a linear device. If the signals $s(n)$ and/or $x_U(n)$ were to have biases in them, the system would not be linear.

(7) **Decimation** of the $s(n)$ and $x_U(n)$ signals to remove out-of-passband zeros from the frequency spectra. This helps us cope with the fact that the problem is ill-posed and ill-conditioned. Removing the out-of-passband spectral zeros makes the problem well-conditioned.

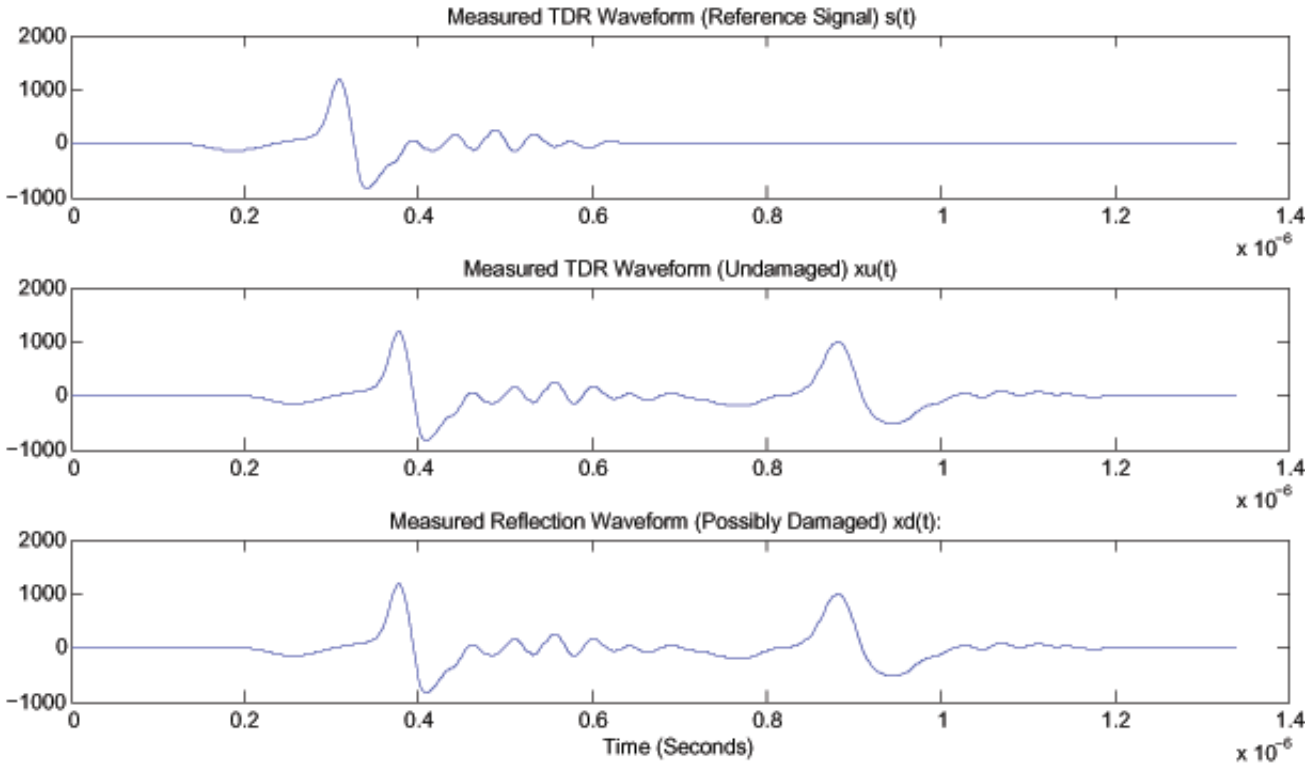


Figure 6.1: Three cut signals: (Top) A reference signal $s(n)$ that has been cut from a longer raw signal, (Middle) A measured signal from an undamaged cable that has been cut from a full-length raw signal, and (Bottom) A measured signal from a damaged cable that has been cut from a full-length raw signal.

6.1.1 Creating a Reference Signal

A key component of the system identification approach is providing a good quality reference signal [15]. A reference signal can be created using one of two basic methods.

(1) Use a reference signal cut from a signal measured for very long cable, if one is available. This is the preferred method, because the reference signal will not contain the unwanted reflections present in measurements from a shorter cable. For some of our work, we were able to obtain a long cable (see the chapters on experiments). These long cables are very similar to the cables of interest, but they are not exactly the same. For our purposes, however, we found that the reference signals obtained from them were very useful in system identification and provided superior results to the reference signals cut from signals derived from regular measurements as described next.

(2) Cut a reference signal from a signal measured from a regular cable of the same length as the cables of interest. This method is not preferred and is used only when a long cable is not available. Multiple reflections can occur before the reflection from the front of the cable is finished ringing. This makes it much more difficult to identify a useful model.

The MATLAB pre-processing codes discussed in the chapter on software give the user the ability to cut out a piece of a given waveform, and zero-pad it on both ends for use as a reference signal. The user must provide a set of parameters such as the time at which to start cutting, number of seconds of data to save after cutting, desired length of the reference signal, etc.

6.1.2 Temporal Shifting/Alignment of the Signals

Figure (6.2) displays a cut and post-padded reference signal $s(n)$ in red plotted over the measurement $x_U(n)$ in blue for delay/alignment comparison. The reference signal was taken from a long cable as described in the last section.

6.1.3 Removing the Mean and/or Trend from the Signals

Removing any statistical mean and/or trend from the data is done to help satisfy the linearity assumption for the system under test. A system with any form of bias (such as a mean and/or trend in the measurements) cannot, by definition, be linear [60, 61]. Since the system identification algorithm assumes linearity, it is very important to execute this step. Generally, any bias in the measurements is an artifact of the measurement process, and should be removed in any case. Figure (6.3) shows the signals $s(n)$, $x_U(n)$ and $x_D(n)$ from Figure (6.1) after having their means and trends removed.

6.1.4 Decimation

To decimate a signal is to reduce its sampling frequency f_s. To interpolate a signal is to increase its sampling frequency. When taken together, Decimation and Interpolation are commonly called Resampling [15, 16, 36, 37, 55]. The author has published in the area of resampling signals and the importance of resampling to the impulse estimation problem [16, 36, 37]. The signals measured in the NDE lab are usually oversampled, in the sense that the sampling frequency is much larger than twice the bandwidth of the signal spectrum (the Nyquist frequency). This is good practice, because it ensures that the data are not aliased [61]. However, for impulse response estimation, it has a very negative effect that can be remedied easily by decimating the data to reduce the available digital frequency bandwidth so it approximates the support of the spectrum of $s(n)$.

For our data set, the sampling frequency is $f_s = 4.096e8\ Hz$, so the folding frequency is $f_s/2 = 2.048e8\ Hz$. We see by visual inspection of Figure (6.4) that the bandwidth is approximately 3.4e8 Hz. Note that due to

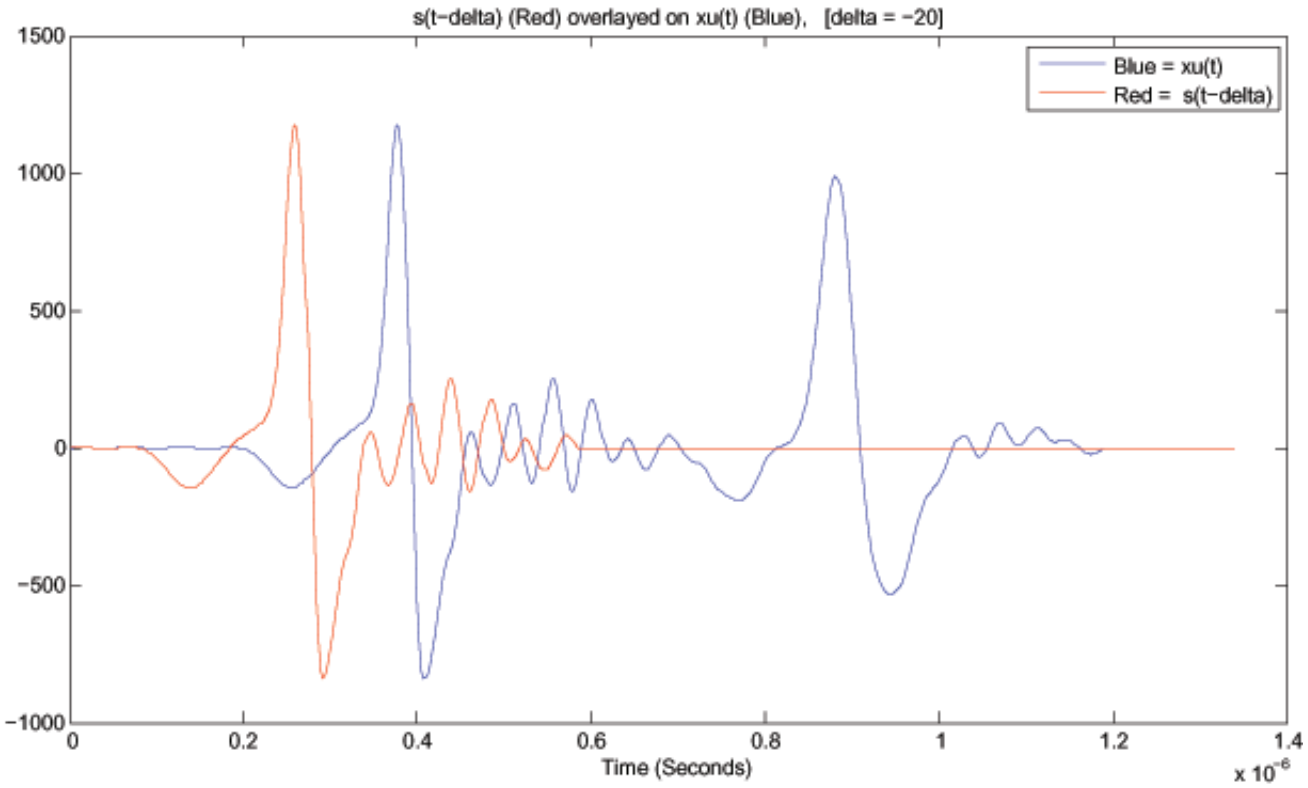


Figure 6.2: A reference signal $s(n-\Delta)$ cut from a signal measured for an extra long cable is displayed in red. The reference signal is plotted over the signal $x_U(n)$ for comparison and to show that the delay Δ has ensured causality for the cable system. Note that both signals have been zero-padded to ensure that they have the same length.

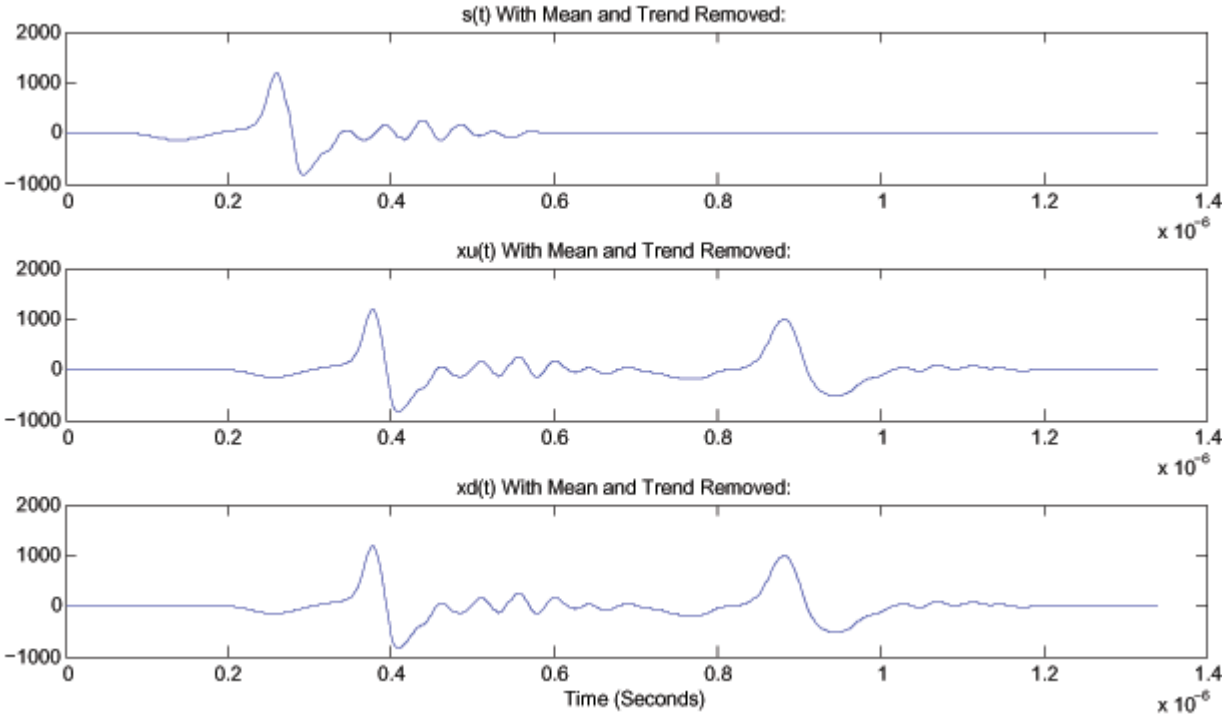


Figure 6.3: The figure shows the signals $s(n)$, $x_U(n)$ and $x_D(n)$ from Figure (6.1) after their means and trends have been removed.

oversampling, the number of spectral zeros is very large, making R_{ss}, the autocorrelation matrix of the reference signal, ill-conditioned. Decimation can remove the high-frequency zeros, but the low frequency zeros and any in-band zeros must be dealt with in other ways. Methods include regularization and applying bandlimited spectrum extrapolation algorithms as described in [15].

An example of decimating the signals is depicted in the figures. Figures (6.4), (6.4) and (6.6) show $|S(f)|^2$, $|X_U(f)|^2$ and $|X_D(f)|^2$ before decimation. The signals are clearly oversampled, as the folding frequency is much larger than the bandwidth.

Figures (6.7), (6.8) and (6.9) show $|X(f)|^2$ and $|Y(f)|^2$ after decimation by a factor of $M = 6$. Note that the high frequency spectral zeros have been effectively eliminated.

The final decimated signals $s(n)$, $x_U(n)$ and $x_D(n)$ in the time domain are depicted in Figure (6.10).

6.2 Signal Pre-Processing for ROC Building

Several preprocessing steps are executed for the System Identification Step. These steps are summarized briefly here.

(1) **Auto-Alignment** of the raw signals to be used for system identification. See the earlier discussion on auto-alignment and the chapter on software tools.

(2) **Cutting** the desired signal out of the raw signal. Here, we wish to use only the first two cable reflections, because the rest are not useful for damage detection. See the discussion on signal cutting in the last section and the chapter on software tools.

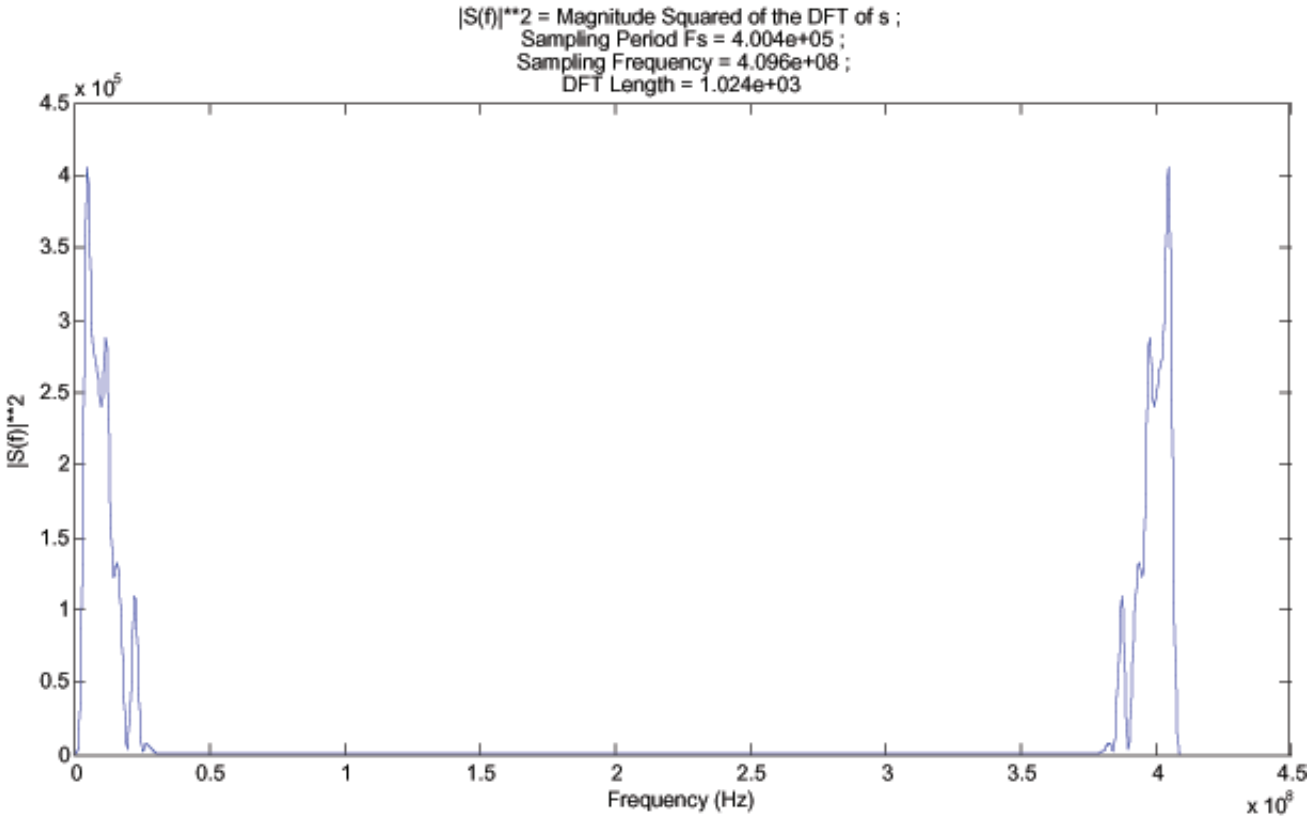


Figure 6.4: The frequency spectrum (DFT) $|S(f)|^2$ before decimation. The sampling frequency is $f_s = 4.096e8\ Hz$, so the folding frequency is $f_s/2 = 2.048e8\ Hz$. We see by visual inspection that the bandwidth is approximately 3.4e7 Hz. The signal is clearly oversampled, as the folding frequency is much larger than the bandwidth.

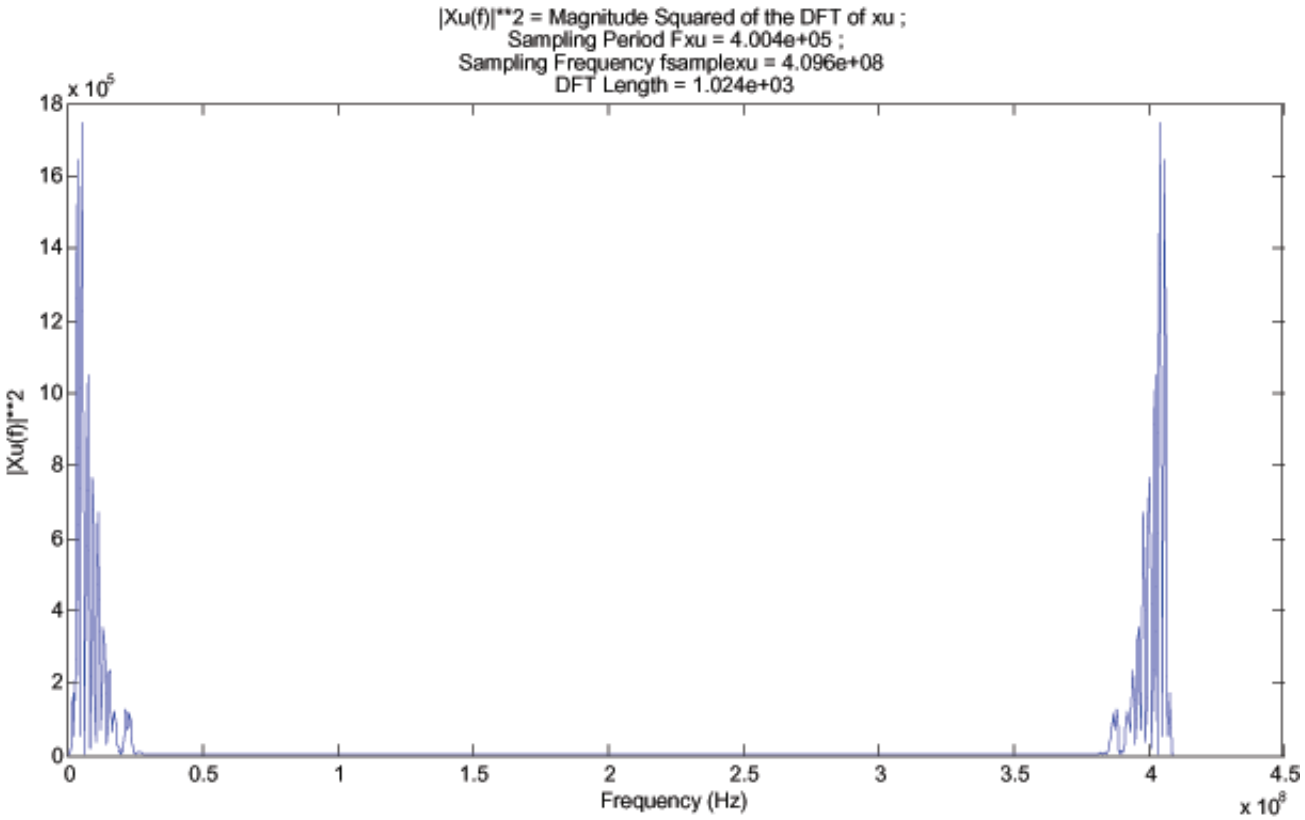


Figure 6.5: The frequency spectrum (DFT) $|X_U(f)|^2$ before decimation. The sampling frequency is $f_s = 4.096e8\ Hz$, so the folding frequency is $f_s/2 = 2.048e8\ Hz$. We see by visual inspection that the bandwidth is approximately 3.4e8 Hz. The signal is clearly oversampled, as the folding frequency is much larger than the bandwidth.

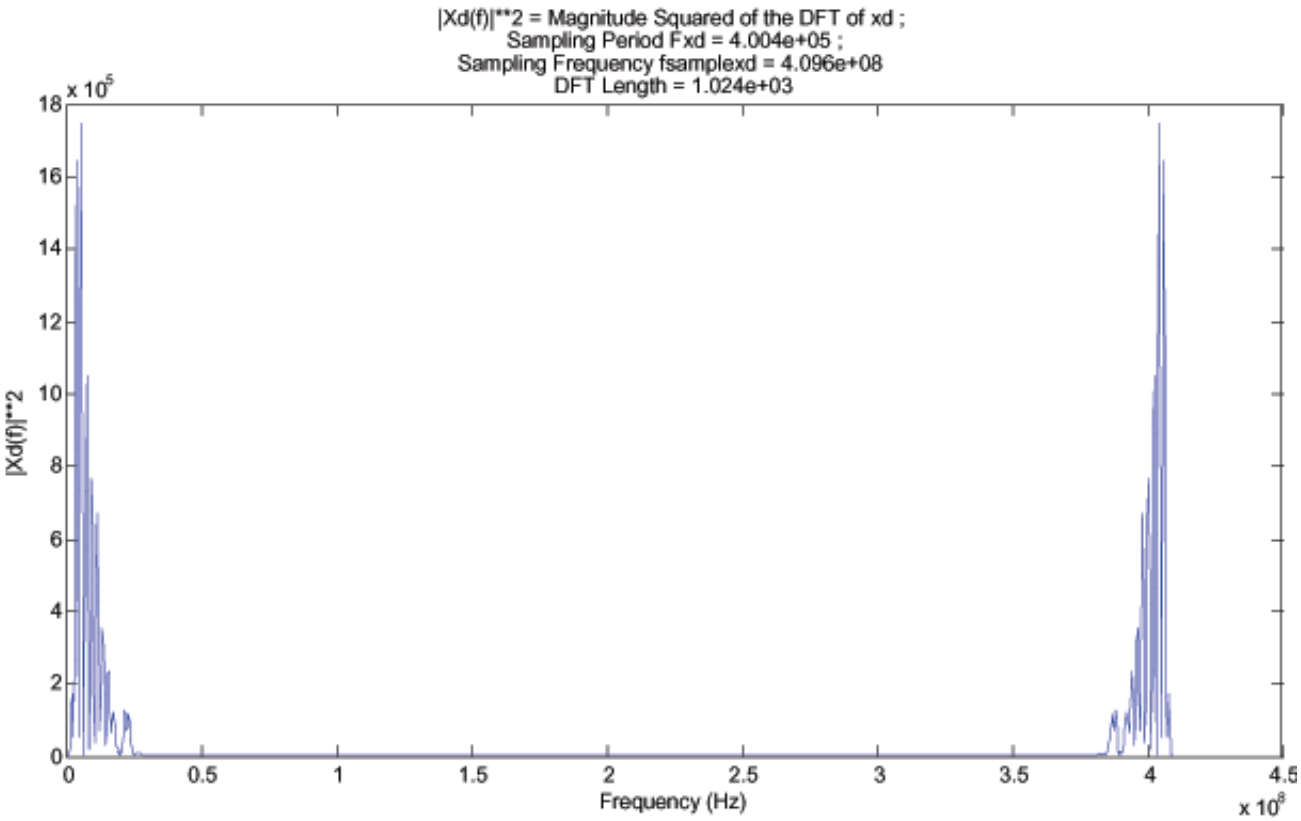


Figure 6.6: The frequency spectrum (DFT) $|X_D(f)|^2$ before decimation. The sampling frequency is $f_s = 4.096e8\ Hz$, so the folding frequency is $f_s/2 = 2.048e8\ Hz$. We see by visual inspection that the bandwidth is approximately 3.4e8 Hz. The signal is clearly oversampled, as the folding frequency is much larger than the bandwidth.

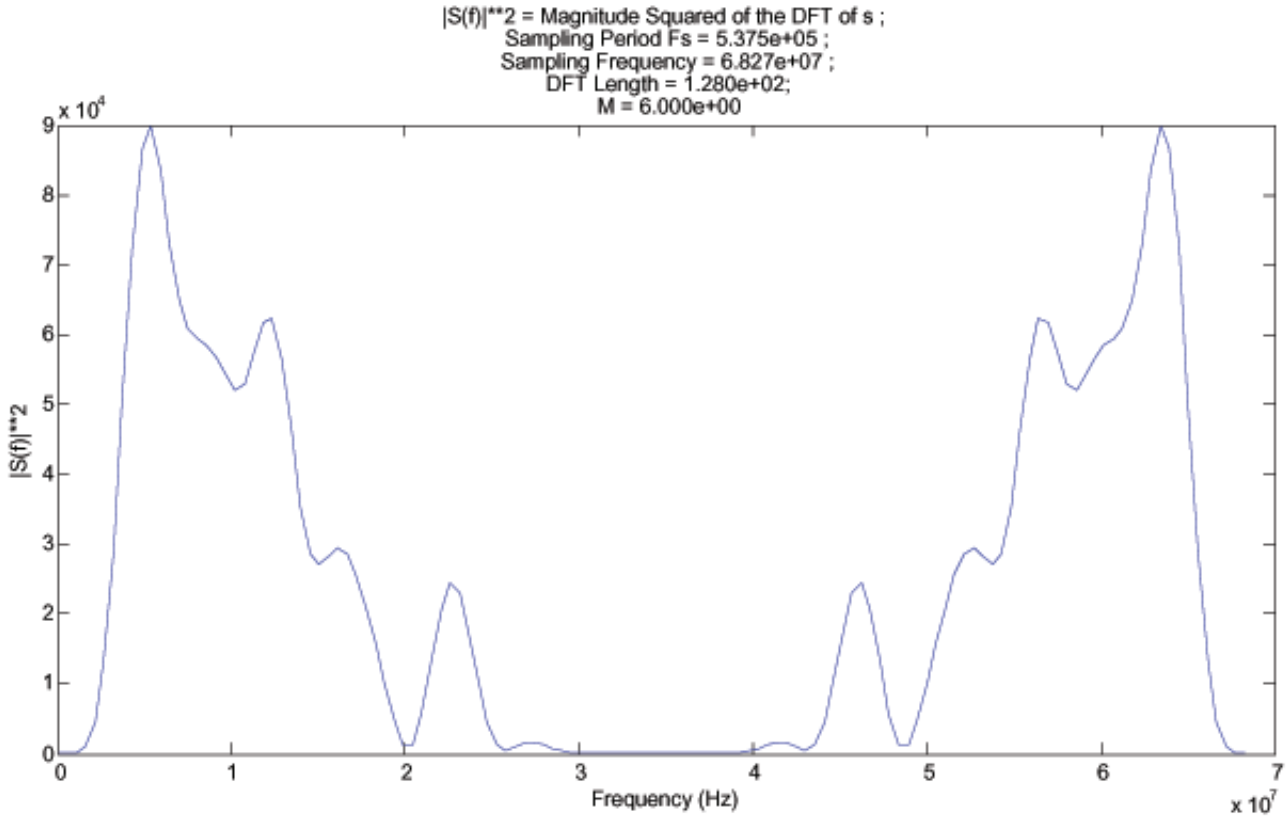


Figure 6.7: The frequency spectrum (DFT) $|S(f)|^2$ after decimation by a factor of 6. The new sampling frequency is $f_s^{'} = 6.827E7\ Hz$, so the folding frequency is $f_s/2 = 3.424E7\ Hz$. We see by visual inspection that the bandwidth is approximately the same as the new folding frequency.

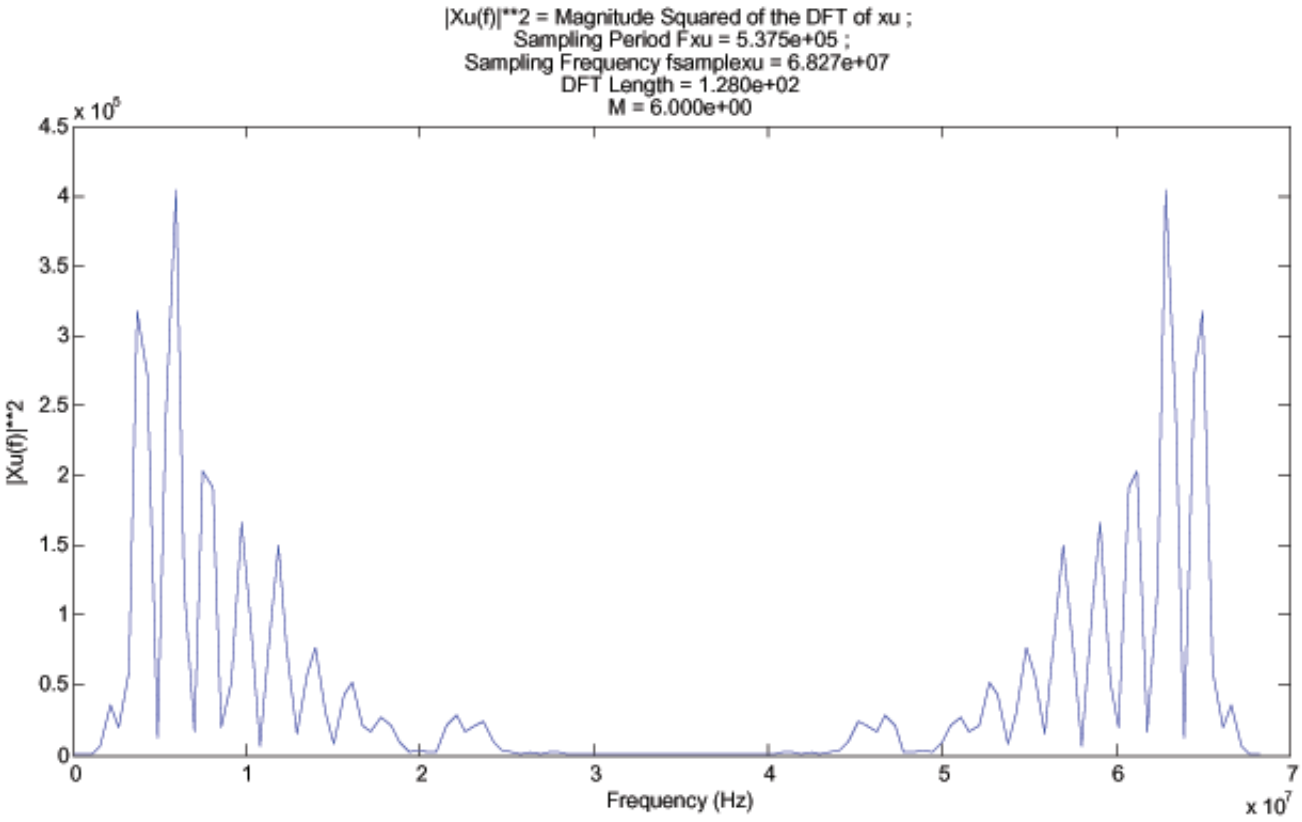


Figure 6.8: The frequency spectrum (DFT) $|X_U(f)|^2$ after decimation by a factor of 6. The new sampling frequency is $f_s^{'} = 6.827E7\ Hz$, so the folding frequency is $f_s/2 = 3.424E7\ Hz$. We see by visual inspection that the bandwidth is approximately the same as the new folding frequency.

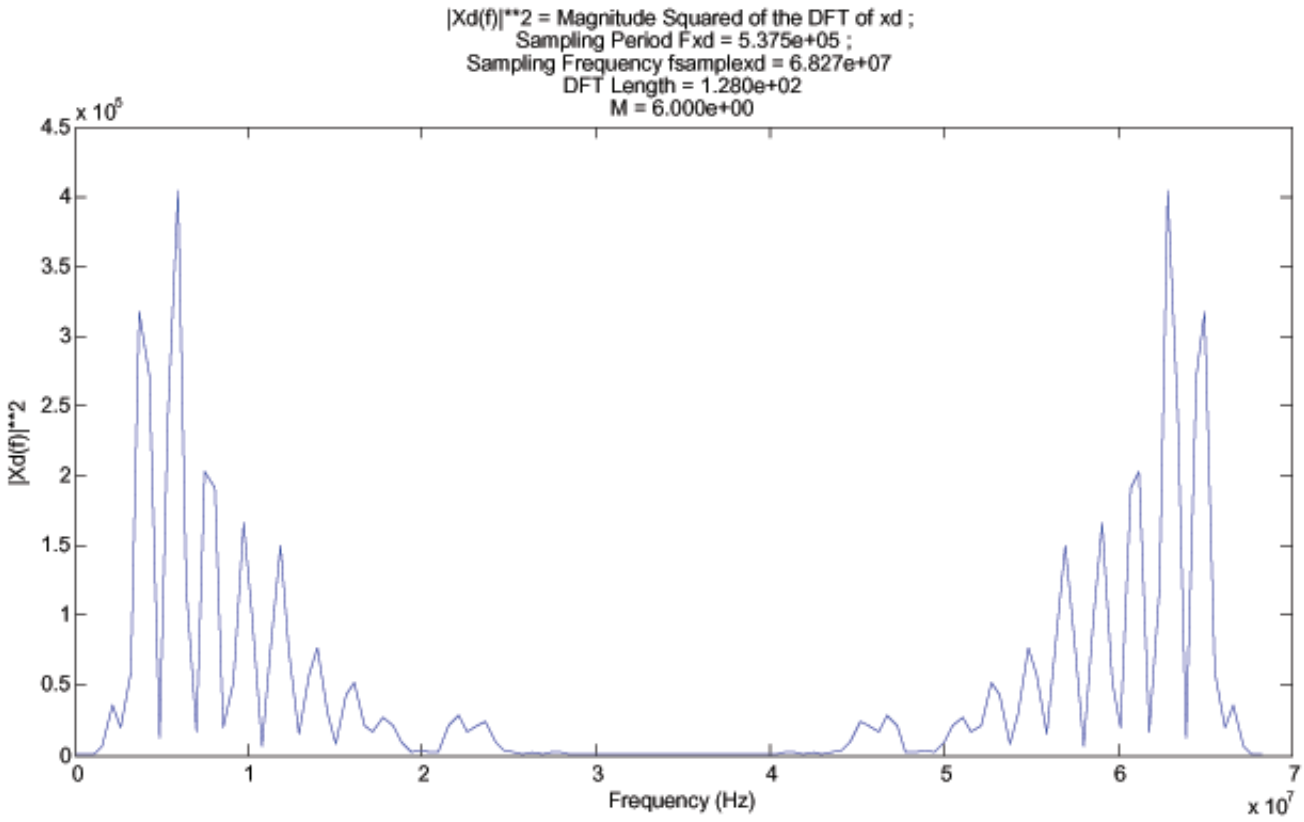


Figure 6.9: The frequency spectrum (DFT) $|X_D(f)|^2$ after decimation by a factor of 6. The new sampling frequency is $f_s^{'} = 6.827E7\ Hz$, so the folding frequency is $f_s/2 = 3.424E7\ Hz$. We see by visual inspection that the bandwidth is approximately the same as the new folding frequency.

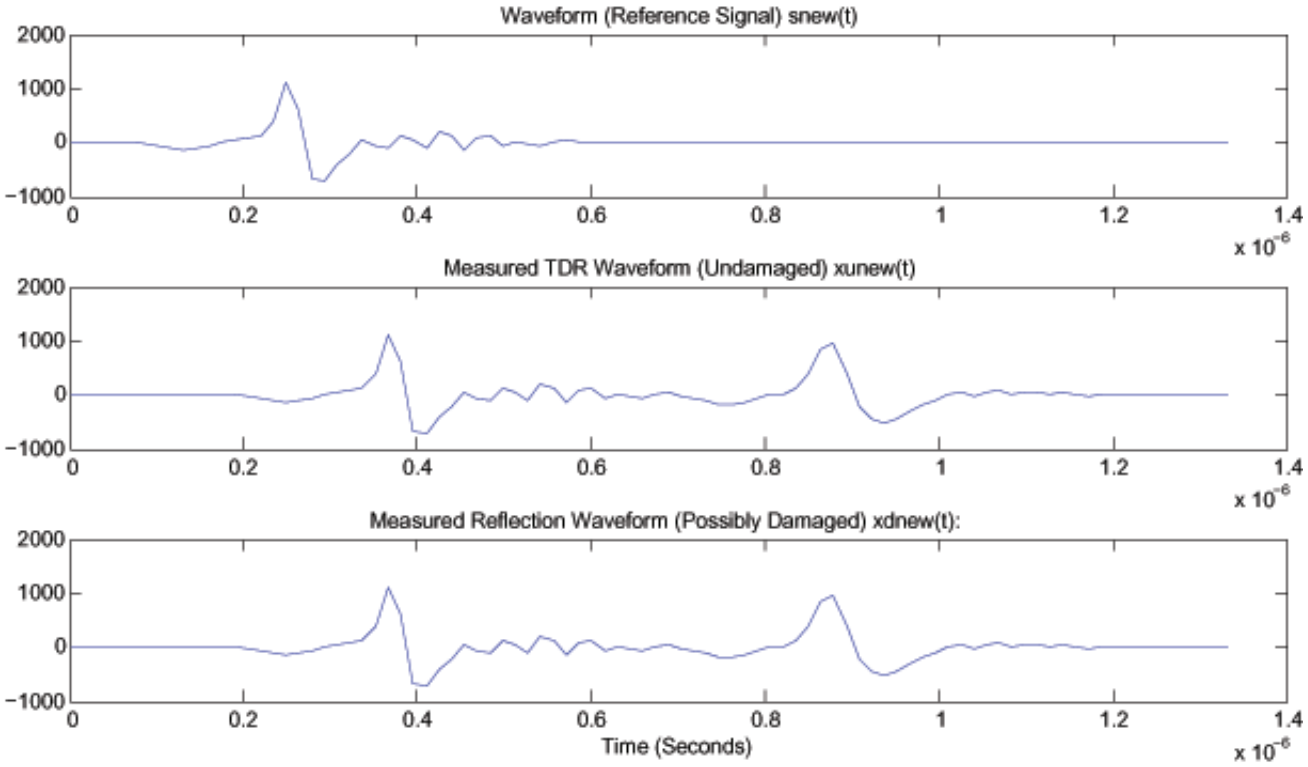


Figure 6.10: The final decimated signals $s(n)$, $x_U(n)$ and $x_D(n)$ in the time domain are plotted for comparison with Figure (6.2). Note that, of course, the time domain plots appear to be less smooth because the sampling frequency has been reduced. However, the information in the signal has been preserved according to the Nyquist criterion.

(3) **Removing the Mean and any Linear Trend** from the signals to ensure that we satisfy the linearity assumption. This must be done because the cable is assumed to be a linear device. If the signals $s(n)$ and/or $x_U(n)$ were to have biases in them, the system would not be linear. See the discussion on mean and trend removal in the last section and the chapter on software tools.

(4) **Decimation** of the $s(n)$ and $x_U(n)$ signals to remove out-of-passband zeros from the frequency spectra. This helps us cope with the fact that the problem is ill-posed and ill-conditioned. Removing the out-of-passband spectral zeros makes the problem well-conditioned. See the discussion on decimation in the last section and the chapter on software tools.

Chapter 7

Detection Performance Measurement Algorithms

We measure the performance of the detector using the Receiver Operating Characteristic (ROC) Curve from communication theory [50, 51]. The ROC curve is a plot of probability of detection vs. probability of false alarm. A statistical confidence interval should be calculated about the probability of correct classification for each point in the ROC curve, using the techniques in the next section.

Using general notation, we consider the two-class hypothesis testing problem in which the goal is to decide between two hypotheses, H_0 and H_1, as depicted in Figure (7.5). For the cable damage detection problem, we let H_0 denote the null hypothesis, or the hypothesis in which the cable is not damaged. We let H_1 denote the hypothesis in which the cable is damaged. For other hypothesis testing procedures, we use different definitions. For example for whiteness testing, H_0 denotes the hypothesis in which the innovations are white and H_1 denotes the hypothesis in which the innovations are not white.

The cable and signal ensembles used for ROC curve generation are discussed in the following sections and are summarized in Figure (7.1), Figure (7.2), and Figure (7.7).

7.1 Ensemble of Cables and Signals for ROC Curve Generation

Building a ROC requires having an ensemble of statistical samples for hypothesis H_0 and corresponding ensemble of statistical samples for hypothesis H_1. For the cable damage detection problem, the ensembles consist of ensembles of TDR signals for undamaged and damaged cases. Let us use the following notation. The discrete time index for the signals is $n = 1, 2, \ldots, M$. The index on the cable number is $k = 1, 2, \ldots, K$. We assume that for each cable, we measure a small ensemble of J signals, inserting the cable into the test fixture and connected it to the PIU once per measurement. Let the index on these signals be $j = 1, 2, \ldots, J$. For our experiments, we used $J = 3$.

7.1.1 Organization of Cable Sets as in Figure (7.1)

Organization of Cable Sets: The sets of cables are organized according to this block diagram. Given K_U undamaged cables (blue box at the top), we define several cable subsets. We set aside K_{Uc} cables for contingencies, K_D cables to be damaged with various categories of damage, and K_E cables for experimentation. We allow K_E cables for use in trying various types of cable insults and any other experiments that we may want to do. This is a set for

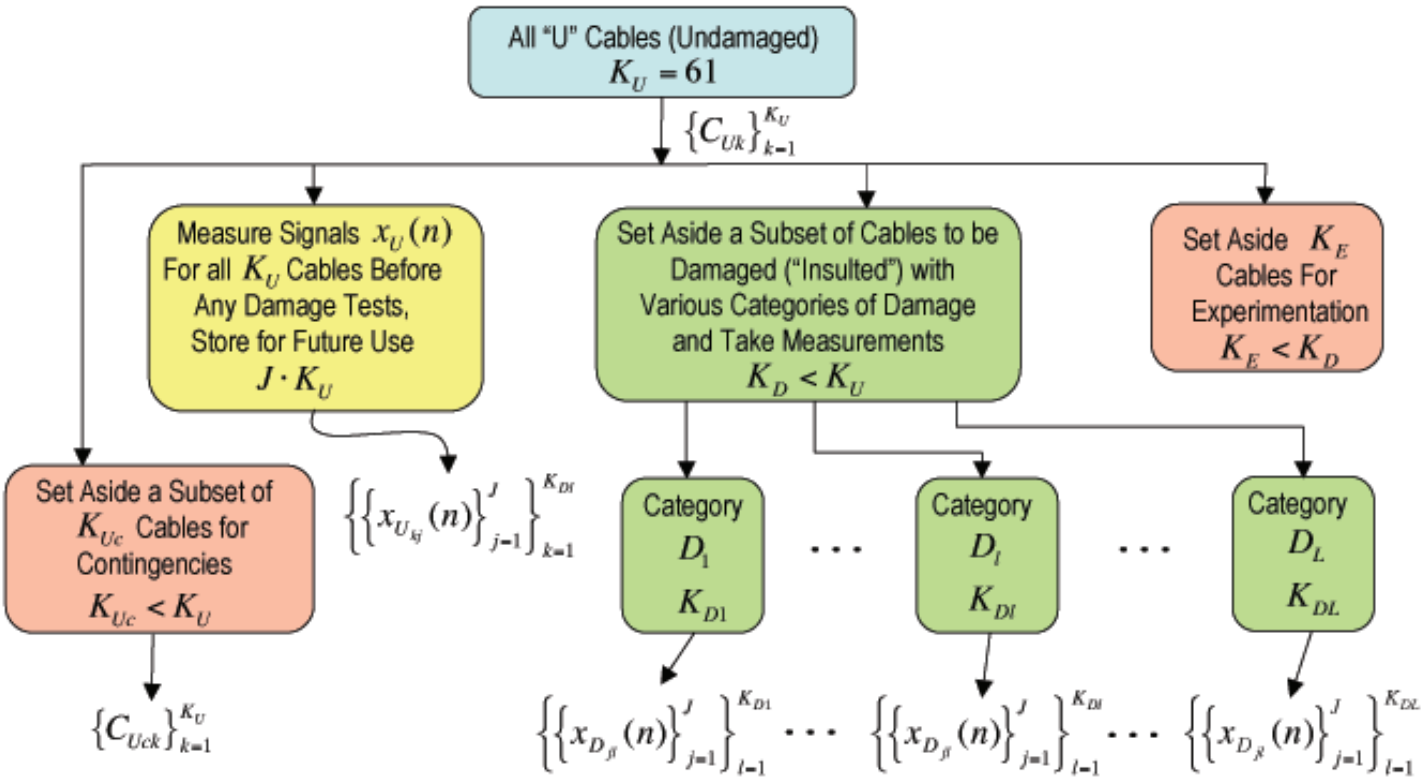


Figure 7.1: Organization of Cable Sets: The sets of cables are organized according to this block diagram.

contingencies. Note that before damaging any cables, we measure and store J baseline signals for each cable, for a total of JK_U baseline signals. We divide the damaged cables into sets with L damage types, as shown in the figure.

7.1.2 Organization of Cable Signal Sets as in Figure (7.2)

Organization of Cable Signal Sets: We organize sets of cables to be used for System Identification, ROC Building and Testing of a Given Cable. For System ID, we create a set of P undamaged cables for creating the reference signal $s(n)$, and a set of P cables for creating the $x_U(n)$ signal. For each cable, we measure J signals to be ensemble averaged to form the signals $\bar{s}(n)$ and $\bar{x}_U(n)$. Thus, there are $Q = JP$ signals for each subset. For ROC building, we use a subsets of undamaged and damaged cables. We use $Q^{'} \leqslant Q$ of the undamaged cables, where $Q = JK_{Dl}$ as shown in the figure. We use K_{Dl} damaged cables from damage category l. The ROC is built from these sets. For testing a single cable, the signal $\widehat{x}_U(n)$ and the optimal WSSR window length W^* are passed to the WSSR test algorithm to produce the detection decision.

7.2 The Bayes Decision Criterion

This section summarizes the concepts involved in constructing a detector/classifier for the binary hypothesis case [50, 51, 41, 42]. Consider the decision problem in which a source of some kind generates an output that consists of one of two choices corresponding to two hypotheses H_0 and H_1. Each hypothesis maps to a point in observation space. The observation space corresponds to a set of N observations that can be denoted by the observation vector

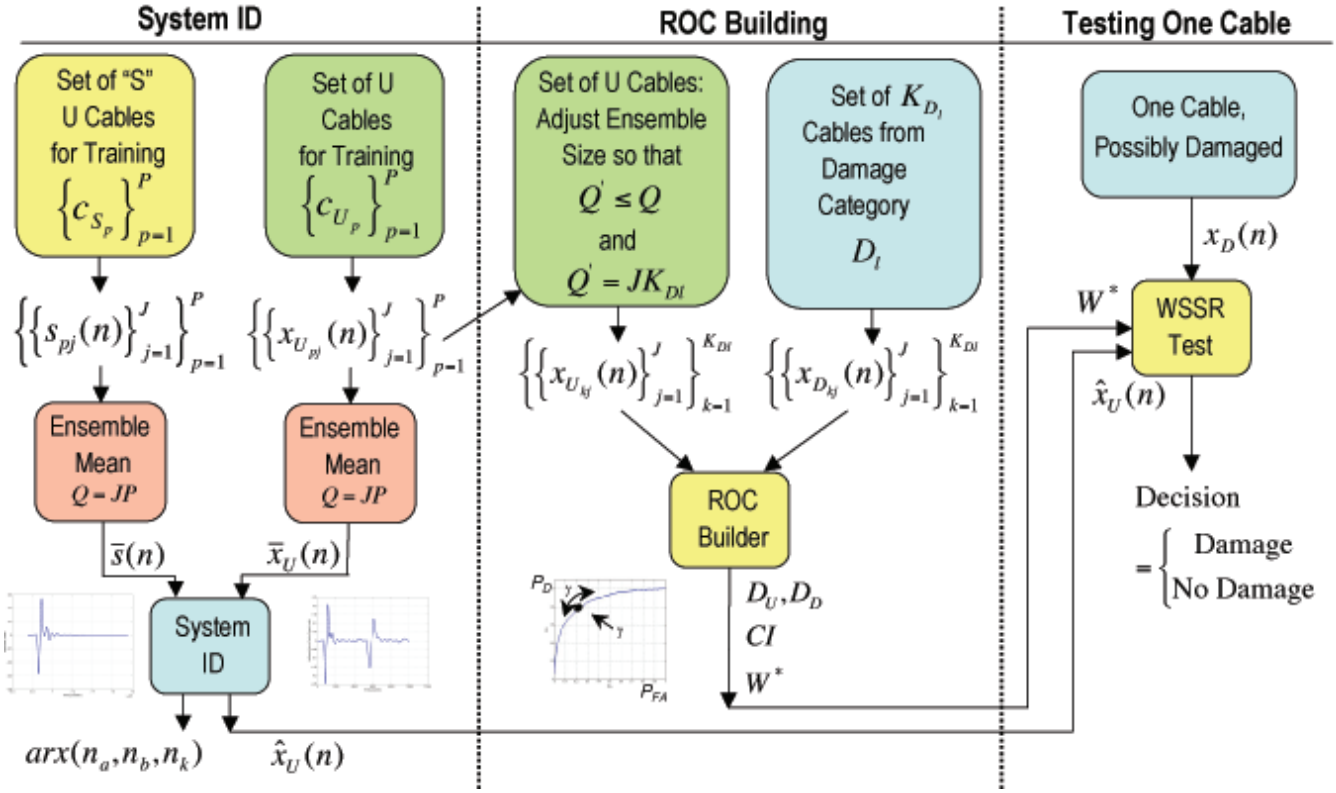


Figure 7.2: Organization of Cable Signal Sets: We organize sets of cables to be used for System Identification, ROC Building and Testing of a Given Cable.

X:

$$\underline{X} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} \qquad \text{(Observation Vector)} \tag{7.1}$$

The system generates observations according to two conditional probability densities $f(\underline{X}|H_0)$ and $f(\underline{X}|H_1)$. We know that either H_0 or H_1 is true and we are required to make a choice between them. Each time the experiment is conducted, one of four events can occur: (1) H_0 is true and we declare H_0, (2) H_0 is true and we declare H_1, (3) H_1 is true and we declare H_1, (4) H_1 is true and we declare H_0. The first and third alternatives correspond to correct choices. The second and fourth alternatives correspond to errors. The purpose of a decision criterion is to assign relative importance to the four possibilities.

The Bayes test assumes that there exist prior probabilities (priors) for the hypotheses and costs associated with the four courses of action. The priors $P(H_0)$ and $P(H_1)$ represent information available about the source prior to conducting the experiments. The costs for the four possible courses of action are given by C_{00}, C_{10}, C_{11} and C_{01}, where C_{ij} is the cost of deciding H_i given that H_j is true. Once the costs have been assigned, the decision rule is based on minimizing the expected cost, which is known as the Bayes risk $\mathfrak{R}$:

$$\mathfrak{R} = \sum_{i=0}^{1} \sum_{j=0}^{1} C_{ij} P(H_i|H_j) P(H_j) \qquad \text{(Bayes Risk)} \tag{7.2}$$

We assume throughout this discussion that the cost of an incorrect decision is higher than the cost of a correct decision. In other words, $C_{10} > C_{00}$ and $C_{01} > C_{11}$. Under this assumption, the detector that minimizes the Bayes risk is given by the following:

$$\frac{f(\underline{X}|H_1)}{f(\underline{X}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \frac{P(H_0)(C_{10} - C_{00})}{P(H_1)(C_{01} - C_{11})} \qquad \text{(Bayes Decision Criterion)} \tag{7.3}$$

The ratio of the conditional densities is called the likelihood ratio and is denoted by $\Lambda(\underline{X})$:

$$\Lambda(\underline{X}) = \frac{f(\underline{X}|H_1)}{f(\underline{X}|H_0)} \qquad \text{(Likelihood Ratio)} \tag{7.4}$$

Because this is a ratio of two functions of a random variable, the likelihood ratio is a random variable. A very important result is that regardless of the dimensionality of the observations X, the likelihood ratio $\Lambda(\underline{X})$ is a one-dimensional variable. This idea is of fundamental importance in hypothesis testing. Regardless of the dimension of the observation space, the decision space is one-dimensional. The quantity on the right hand side of the relation 7.3 is the threshold of the test and is denoted by η:

$$\eta \triangleq \frac{P(H_0)(C_{10} - C_{00})}{P(H_1)(C_{01} - C_{11})} \qquad \text{(Threshold)} \tag{7.5}$$

Thus, the Bayes criterion leads to a likelihood ratio test:

$$\Lambda(\underline{X}) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \qquad \text{(Likelihood Ratio Test)} \tag{7.6}$$

We see that the test threshold allows for weighting according to the priors and the costs. This allows the user flexibility in choosing a threshold that is best for the problem at hand.

Another fundamental concept is that of a sufficient statistic $l(\underline{X})$. This is a function of the observed data which has the property that $\Lambda(\underline{X})$ can be written as a function of $l(\underline{X})$. This means that when making a decision, knowing the value of the sufficient statistic is just as good as knowing X.

We can now generate a receiver operating characteristic (ROC) curve that depicts the tradeoff between probability of detection and probability of false alarm (see Figure (7.5) and (7.6). Details of ROC curve development are given in the next section. Let us now consider some important special cases of the Bayes decision rule.

7.2.1 Bayes Risk Special Case When $C_{00} = C_{11} = 0$ and $C_{01} = C_{10} = 1$

An important special case of the Bayes criterion is that in which a correct classification is assigned zero cost and an incorrect classification is full cost. In this case, we assign $C_{00} = C_{11} = 0$ and $C_{01} = C_{10} = 1$. Inserting these values in the Bayes Risk of Equation (7.2), we obtain:

$$\begin{aligned} \mathfrak{R} &= P(\text{Error}) = P(H_0, H_1) + P(H_1, H_0) && (7.7)\\ &= P(H_0|H_1)P(H_1) + P(H_1|H_0)P(H_0) && (7.8) \end{aligned}$$

This version of the Bayes risk is called probability of error, and is a very effective criterion for evaluating detector/classifier performance [50, 51, 41, 42]. Alternatively, we often use the fact that $P(\text{Error})+P(\text{Correct Classification}) = 1$ and define the probability of correct classification:

$$\begin{aligned} P(\text{Correct Classification}) = P(CC) &= P(H_1, H_1) + P(H_0, H_0) && (7.9)\\ &= P(H_1|H_1)P(H_1) + P(H_0|H_0)P(H_0) && (7.10) \end{aligned}$$

In this work, we use probability of correct classification as a very useful performance measure.

7.2.2 Probability of Correct Classification Special Case When Priors are Equal

Often in practice there exists insufficient information about an experiment to allow assignment of prior probabilities P_0 and P_1. In this case, it is common to assume that the priors are equal, so $P_0 = P_1 = 1/2$. Under this condition, the probability of correct classification becomes

$$P(CC) = \frac{1}{2}[P(H_1|H_1) + P(H_0|H_0)] \tag{7.11}$$

This can now be written in terms of the probability of detection and probability of false alarm as follows:

$$P(CC) = \frac{1}{2}[P_D + (1 - P_{FA})] \tag{7.12}$$

7.2.3 The Bayes Minimax Test

A Bayes test designed to minimize the maximum possible risk is called a minimax test. For any choice of decision regions the risk expression can be written in terms of the quantities in Figures (7.5) and (7.6) as follows:

$$\mathfrak{R} = P(H_0)C_{10} + P(H_1)C_{11} + P(H_1)(C_{01} - C_{11})P_M - P(H_0)(C_{10} - C_{00})(1 - P_{FA}) \tag{7.13}$$

where $P_M = 1 - P_D$ is the probability of a miss. To minimize the maximum risk, we use a Bayes test designed assuming $P(H_1) = 1$. This implies that the coefficient of $P(H_1)$ must be zero, and the resulting minimax equation is:

$$(C_{11} - C_{00}) + (C_{01} - C_{11})P_M - (C_{10} - C_{00})P_{FA} = 0 \tag{7.14}$$

Assuming that $C_{00} = C_{11} = 0$ and substituting $P_M = 1 - P_D$, the risk becomes

$$\mathfrak{R} = P(H_0)C_{10}P_{FA} + P(H_1)C_{01}(1 - P_D) \tag{7.15}$$

and the minimax equation is

$$C_{01}(1 - P_D) = C_{10}P_{FA} \tag{7.16}$$

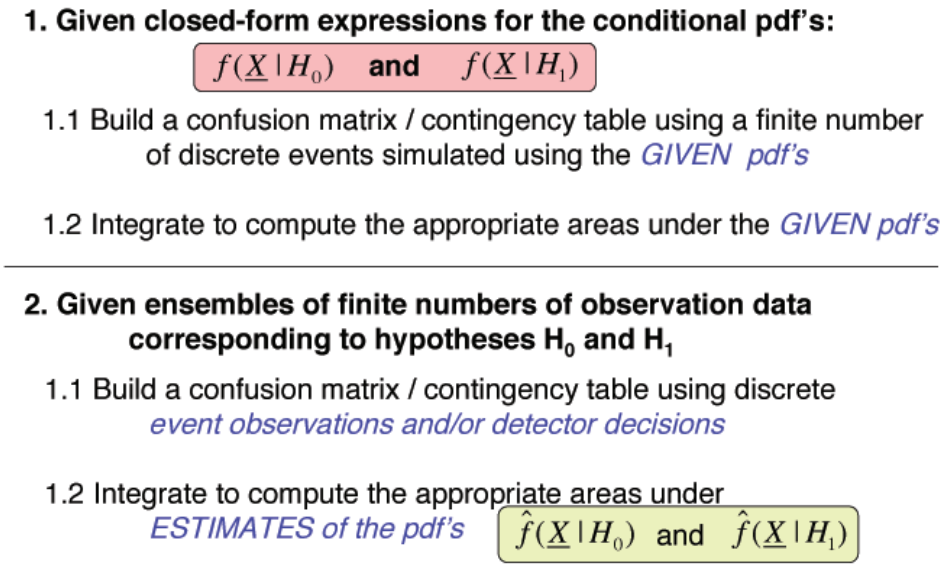


Figure 7.3: There are two general ways in which a ROC Curve can be constructed. If one is given closed-form expressions of the conditional probability density functions, then one can integrate them as in Figure (7.5) and (7.6). If one is given ensembles of finite numbers of observation data corresponding to the hypotheses, then a contingency table can be constructed as in Figure (7.4). The probability of detection and probability of false alarm are taken directly from the contingency table. Or, if estimates of the pdf's are available, they can be integrated to obtain the probabilities, as in Figure (7.5) and (7.6).

7.3 Receiver Operating Characteristic Curve

In this discussion on ROC curve generation, let us use the following notation. The discrete time index for the signals is $n = 1, 2, \ldots, M$. The index on the WSSR window length is $j = 1, 2, \ldots, J$. The index on the signal number is $k = 1, 2, \ldots, K$.

Using general notation, we consider the two-class hypothesis testing problem in which the goal is to decide between two hypotheses, H_0 and H_1, as depicted in Figure (7.5). The problem depicted is that of detecting a signal in stochastic noise. The algorithm chosen determines the detection statistic r on which we place a variable threshold r_0. When the threshold is varied across the values of r, a curve is traced out with a shape as shown in Figure (7.5). Families of ROC curves can be formed by varying the signal-to-noise ratio (SNR) or other appropriate quantities of interest.

7.3.1 The Contingency Table or Confusion Matrix

In this cable damage detection problem, the conditional probability density functions are not available, so we use the confusion matrix approach to calculate the ROC curve.

As depicted in Figure (7.2), the results of the WSSR test consist of vectors indicating the binary "truth "and

Truth / Decision	θ_1	θ_2
θ_1	$P(\theta_1 \mid \theta_1) = P(Detection)$ $= \frac{No.\ Samples\ Classified\ \theta_1}{No.\ \theta_1\ Samples}$	$P(\theta_1 \mid \theta_2) = P(\textbf{False Alarm})$ $= \frac{No.\ Samples\ Classified\ \theta_1}{No.\ \theta_2\ Samples}$
θ_2	$P(\theta_2 \mid \theta_1) = P(\textbf{Miss})$ $= \frac{No.\ Samples\ Classified\ \theta_2}{No.\ \theta_1\ Samples}$	$P(\theta_2 \mid \theta_2) = Specificity$ $= \frac{No.\ Samples\ Classified\ \theta_2}{No.\ \theta_2\ Samples}$

$$P(\theta_1 \mid \theta_1) + P(\theta_2 \mid \theta_1) = 1$$
$$P(\theta_1 \mid \theta_2) + P(\theta_2 \mid \theta_2) = 1$$
$$P(Correct\ Classification) = P(CC) = P(\theta_1 \mid \theta_1)P(\theta_1) + P(\theta_2 \mid \theta_2)P(\theta_2)$$

Figure 7.4: Confusion Matrix (Contingency Table) in Theta notation. In this figure, the hypotheses are denoted with the symbol θ rather than the symbol H that is used in the text. Here, θ_1 corresponds to H_1 and θ_2 corresponds to H_0

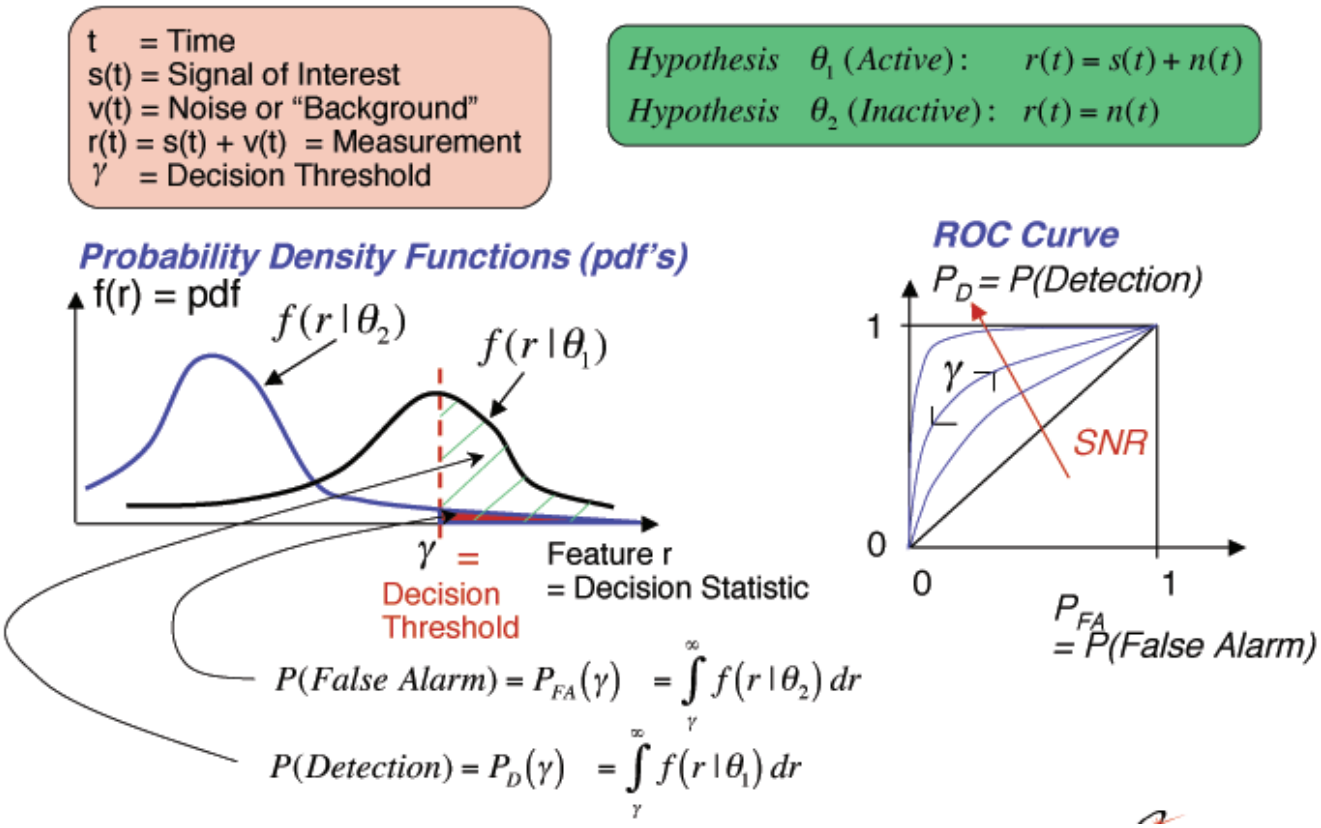
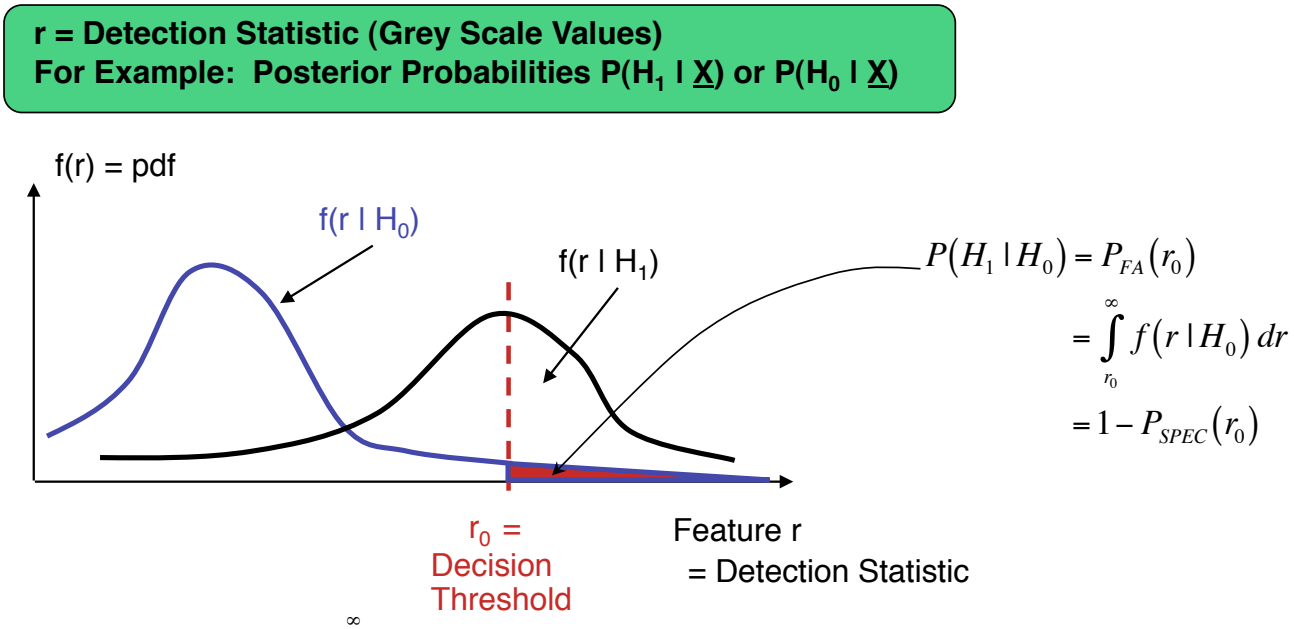


Figure 7.5: The ROC curve can be generated by integrating under the conditional probability density functions to compute P_D and P_{FA}. As the decision threshold is varied, a curve is mapped out in the ROC curve. Families of curves can be developed by letting the signal-to-noise (SNR) ratio vary.



$$P(H_1 \mid H_1) = P_D(r_0) = \int_{r_0}^{\infty} f(r \mid H_1)\, dr = 1 - P(H_0 \mid H_1) = 1 - P_{MISS}(r_0)$$

$$P(H_0 \mid H_1) = P_{MISS}(r_0) = \int_{-\infty}^{r_0} f(r \mid H_1)\, dr = 1 - P(H_1 \mid H_1) = 1 - P_D(r_0)$$

$$P(H_0 \mid H_0) = P_{SPEC}(r_0) = \int_{-\infty}^{r_0} f(r \mid H_0)\, dr$$

Figure 7.6: The details of the integrations under the ROC curve are depicted in this figure. Let r denote the test statistic of interest. Then the various conditional probabilities in the contingency table are computed by the integrations shown.

the binary test declaration. The $J \times 1$ truth vector for the undamaged case is ($k = 1, 2, \ldots, K$):

$$\underline{t}_{Uk} = \begin{bmatrix} t_{Uk}(1) \\ t_{Uk}(2) \\ \vdots \\ t_{Uk}(J) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix} \qquad (J \times 1 \text{ Vector of Zeros}) \tag{7.17}$$

The $J \times 1$ truth vector for the damaged case is ($k = 1, 2, \ldots, K$):

$$\underline{t}_{Dk} = \begin{bmatrix} t_{Dk}(1) \\ t_{Dk}(2) \\ \vdots \\ t_{Dk}(J) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{bmatrix} \qquad (J \times 1 \text{ Vector of Ones}) \tag{7.18}$$

The $J \times 1$ declaration vector for the undamaged case is ($k = 1, 2, \ldots, K$):

$$\underline{d}_{Uk} = \begin{bmatrix} d_{Uk}(1) \\ d_{Uk}(2) \\ \vdots \\ d_{Uk}(J) \end{bmatrix} \qquad (J \times 1) \tag{7.19}$$

The $J \times 1$ declaration vector for the damaged case is ($k = 1, 2, \ldots, K$):

$$\underline{d}_{Dk} = \begin{bmatrix} d_{Dk}(1) \\ d_{Dk}(2) \\ \vdots \\ d_{Dk}(J) \end{bmatrix} \qquad (J \times 1) \tag{7.20}$$

For convenience, we organize the vectors above into four corresponding $J \times K$ matrices that can be used to calculate probability of detection and probability of false alarm as depicted in Figure (7.7). We define:

$$T_U = \begin{bmatrix} 0 \end{bmatrix} \qquad (J \times K \text{ Matrix of Zeros}) \tag{7.21}$$

$$T_D = \begin{bmatrix} 1 \end{bmatrix} \qquad (J \times K \text{ Matrix of Ones}) \tag{7.22}$$

$$D_U = \begin{bmatrix} \underline{d}_{U1}, \underline{d}_{U2}, \ldots \underline{d}_{Uk} \ldots, \underline{d}_{UK} \end{bmatrix} \qquad (J \times K) \tag{7.23}$$

$$D_D = \begin{bmatrix} \underline{d}_{D1}, \underline{d}_{D2}, \ldots \underline{d}_{Dk} \ldots, \underline{d}_{DK} \end{bmatrix} \qquad (J \times K) \tag{7.24}$$

Given the results of the experiment and the matrices above, we can calculate the ROC curve points using the following expressions:

$$P_{FA} = \frac{\text{Number of } H_0 \text{ Signals } e_U(k,n) \text{ Declared } H_1}{\text{Number of } H_0 \text{ Signals } e_U(k,n)} \qquad \text{(Probability of False Alarm)} \tag{7.25}$$

$$P_D = \frac{\text{Number of } H_1 \text{ Signals } e_D(k,n) \text{ Declared } H_1}{\text{Number of } H_1 \text{ Signals } e_D(k,n)} \qquad \text{(Probability of Detection)} \tag{7.26}$$

Because of the matrix organization, the calculations are very simple in the MATLAB language.

7.3.2 Probability of Correct Classification

Recall the definition of probability of correct classification, P_{CC}. For the cable damage detection problem, we let H_0 represent the hypothesis that no damage exists, and let H_1 represent the hypothesis that damage does exist. Recall that the key conditional probabilities are $P(H_1|H_1) = P_D$, the probability of detection (or sensitivity), and $P(H_0|H_0) = P_{spec}$, the specificity. From the contingency tables (confusion matrices) shown earlier, we know that $P_{spec} = 1 - P_{FA}$, where P_{FA} is the probability of false alarm. Exploiting these results, we derive the following expressions:

$$P_{CC} = P(H_1|H_1)P(H_1) + P(H_0|H_0)P(H_0) \tag{7.27}$$

$$= P_D P(H_1) + [1 - P_{FA}]P(H_0) \quad \text{(Probability of Correct Classification)} \tag{7.28}$$

If the application does not allow knowledge of the prior probabilities, a reasonable way to proceed is to assume that the two hypotheses are equally likely, so $P(H_1) = P(H_0) = 1/2$. Under this condition, the expression for P_{CC} is simplified to the following:

$$P_{CC} = \frac{1}{2}\{P_D + [1 - P_{FA}]\} \quad \text{(Probability of Correct Classification)} \tag{7.29}$$

For the cable damage detection problem, some information about the cables is known a priori, so some estimates of the prior probabilities of the two hypotheses could possibly be made. For the purposes of this report, that work was not done because of programmatic constraints on resources. The priors were assumed to be equal. A study of prior probabilities is proposed for future work.

7.3.3 Organization of the ROC Building Algorithm as in Figure (7.7)

Given sets of innovation signals for the undamaged and damaged cases, the WSSR test is calculated to form sets of declaration vectors and truth vectors that are assembled in declaration matrices and truth matrices as shown in the figure. These matrices allow the P_{FA} and P_D calculations that form the ROC curve when the WSSR window length is varied. Recall that the decision threshold is a function of only the window length W.

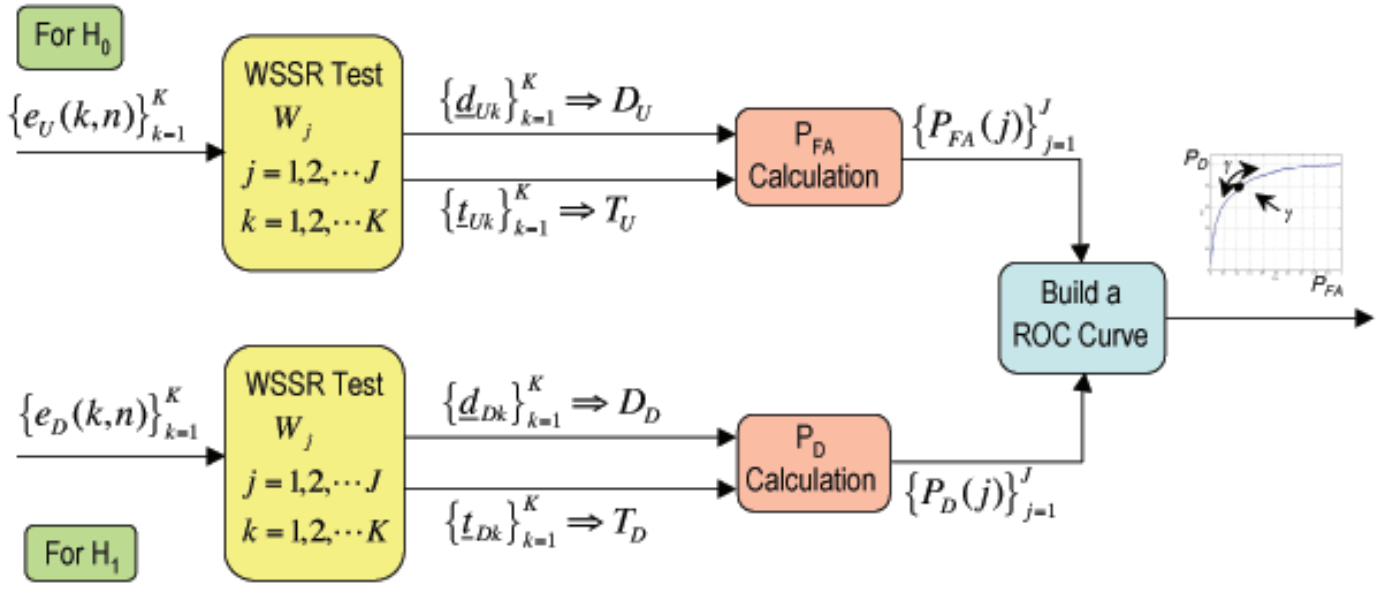
7.4 Statistical Confidence Interval on the Probability of Correct Classification

We are very interested in knowing the confidence with which we can specify the performance of classifiers. In this section, we present algorithms for computing confidence criteria.

In the process of evaluating classification/detector performance, we estimate conditional probabilities based upon experiments with real data and a finite number of statistical samples. We can specify the performance in terms of sensitivity and specificity. In order to specify the performance fully, however, it is desired to specify the confidence we have in the estimates of the conditional probabilities. We can do this by calculating a statistical confidence interval about the probability of correct classification.

The classifier/detector performs a random experiment, the outcome of which can be classified in one of two mutually exclusive and exhaustive ways: success or failure. Success means that the classification is correct. Failure means that the classification is incorrect. Let N equal the number of independent trials. Let $p = P(CC) =$ the probability of success. Assume that p is the same on each repetition. Let $q = 1 - p =$ probability of failure. Now, let X_i be a random variable with $i = 1, 2, \ldots, N$ and

$$X_i = \begin{cases} 0, & \text{if the outcome of the ith trial is failure} \\ 1, & \text{if the outcome of the ith trial is success} \end{cases}$$



$$e_U(k,n) = x_U(k,n) - \hat{x}_U(k,n)$$
$$e_D(k,n) = x_D(k,n) - \hat{x}_U(k,n)$$
$$n = \text{Time Index} = 1,2,\cdots,M$$
$$k = \text{Signal Index} = 1,2,\cdots,K$$
$$j = \text{WSSR Window Index (on } W_j) = 1,2,\cdots,J$$

Figure 7.7: Organization of the ROC building algorithm. Given sets of innovation signals for the undamaged and damaged cases, the WSSR test is calculated to form sets of declaration vectors and truth vectors that are assembled in declaration matrices and truth matrices as shown in the figure. These matrices allow the P_{FA} and P_D calculations that form the ROC curve when the WSSR window length is varied. Recall that the decision threshold is a direct function of only the window length W.

So, we can write $P\{X_i = 1\} = p$ and $P\{X_i = 0\} = 1 - p = q$.

Now, let Y be the sum of successes throughout N repetitions of the experiment.

$$Y = \sum_{i=1}^{N} X_i \tag{7.30}$$

Let $y \in \{y : y = 1, 2, \ldots, N\}$. Then, $Y = y$ iff exactly y of the variables $\{X_1, X_2, \ldots, X_N\}$ have value 1, and the remaining $N - y$ variables equal zero. There are $\binom{N}{y}$ ways in which exactly y ones can be assigned to y of the variables $X_1, X_2, \ldots, X_N$. Since $X_1, X_2, \ldots, X_N$ are mutually stochastically independent, the probability of each of these ways is $p^y(1-p)^{N-y}$. Now, $P(Y = y)$ is the sum of the probabilities of these $\binom{N}{y}$ mutually exclusive events, so

$$P(Y = y) = \begin{cases} \binom{N}{y} p^y (1-p)^{N-y} & , y = 0, 1, 2, \ldots, N \\ 0 & , \text{Elsewhere} \end{cases}$$

This is the probability density function (pdf) of a binomial distributed random variable, so Y is binomial distributed $b(N, p)$. Recall that $\binom{N}{y} = \frac{N!}{y!(n-y)!} = 0$ if $y > N$ or if $y < 0$. Also, note that for a binomial distribution, we have the mean $E\{y\} = Np$, and the variance $\sigma_y^2 = Npq$ [48, 66].

Now for the classification problem in which we conduct an experiment, we can calculate the quantities in the confusion matrix (or contingency table). The maximum likelihood estimate of p is given by:

$$\widehat{p} = \frac{\text{Number of Correct Classifications}}{\text{Number of Test Cases}}$$

We can also write this as follows:

$$\widehat{p} = \frac{y}{N} \tag{7.31}$$

$$\widehat{q} = 1 - \widehat{p} \tag{7.32}$$

We can write the confidence interval about the true value of p as follows, where α is the significance of the test.

$$P\{L < p < U\} = 1 - \alpha \tag{7.33}$$

where L and U are the lower and upper bounds, respectively, of the confidence interval. There are various ways to interpret the meaning of a confidence interval.

(1) This most common interpretation is to read the confidence interval relation above as follows: "With confidence $1 - \alpha$, the true p lies between L and U." However, this interpretation is not generally supported by statistical rigor. The preferred interpretations are given next.

(2) The classical interpretation is to read the confidence interval relation above as follows: "Prior to the repeated independent performances of the random experiment, the probability is $1 - \alpha$ that the random interval (L, U) includes the unknown fixed point (parameter) p [66]."

(3) The frequentist interpretation is to read the confidence interval relation as follows: "The confidence interval is a random interval that covers the true probability with frequency $1 - \alpha$. " This does not mean that a particular interval contains the true value of p with probability $1 - \alpha$ percent. The reason for this somewhat convoluted interpretation is that p is an unknown constant and not a random variable, and we cannot make probability statements about constants within the frequentist framework of statistics.

For the large sample size case ($N \geq 30$), the Gaussian approximation to the distribution of p can be used, and

the 95% confidence interval ($\alpha = .05$) is given by [66]:

$$L = \widehat{p} - 1.96\sqrt{\frac{\widehat{p}\,(1-\widehat{p})}{N}} \qquad \text{and} \tag{7.34}$$

$$U = \widehat{p} + 1.96\sqrt{\frac{\widehat{p}\,(1-\widehat{q})}{N}} \tag{7.35}$$

We can evaluate L and U, and plot them versus p and $\widehat{p}$, for various values of N, as in Figure (7.8). The plot is very instructive in showing how the confidence interval tightens as the sample size increases.

In many applications however, we are limited to small sample sizes, so we are forced to use more accurate estimates of L and U which are valid for small sample sizes. These estimates are given as follows [66]:

$$L = N\widehat{p} + 2 - 2\sqrt{\frac{N\widehat{p}\,(1-\widehat{p})+1}{N+4}} \qquad \text{and} \tag{7.36}$$

$$U = N\widehat{p} + 2 + 2\sqrt{\frac{N\widehat{p}\,(1-\widehat{p})+1}{N+4}} \tag{7.37}$$

For the small sample size case, we can evaluate L and U, and plot them versus p and $\widehat{p}$, for various values of N, as in Figure (7.9). Again, the plot is very instructive in showing how the confidence interval tightens as the sample size increases.

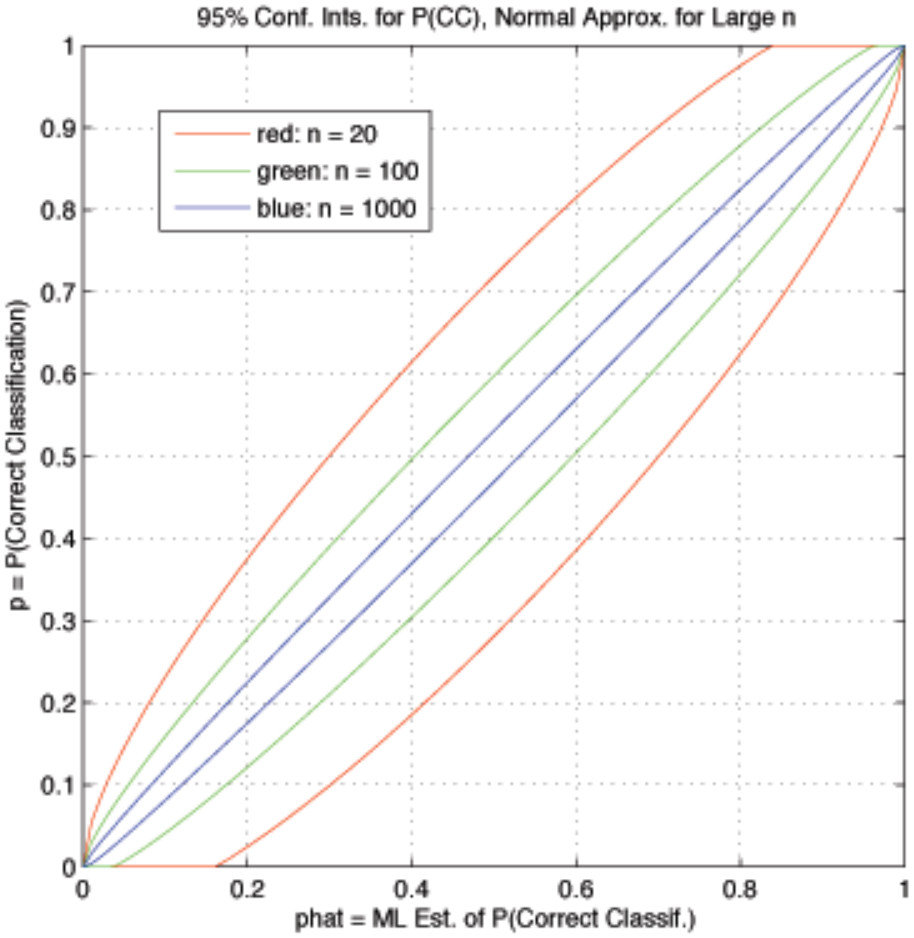


Figure 7.8: Large Sample Size Case: The 95 Percent confidence interval bounds (L, U) for the probability of correct classification are plotted, given the sample size n. Here, we use the Normal approximation ordinarily used for the the case in which n is large. Note how the confidence interval tightens as the sample size increases.

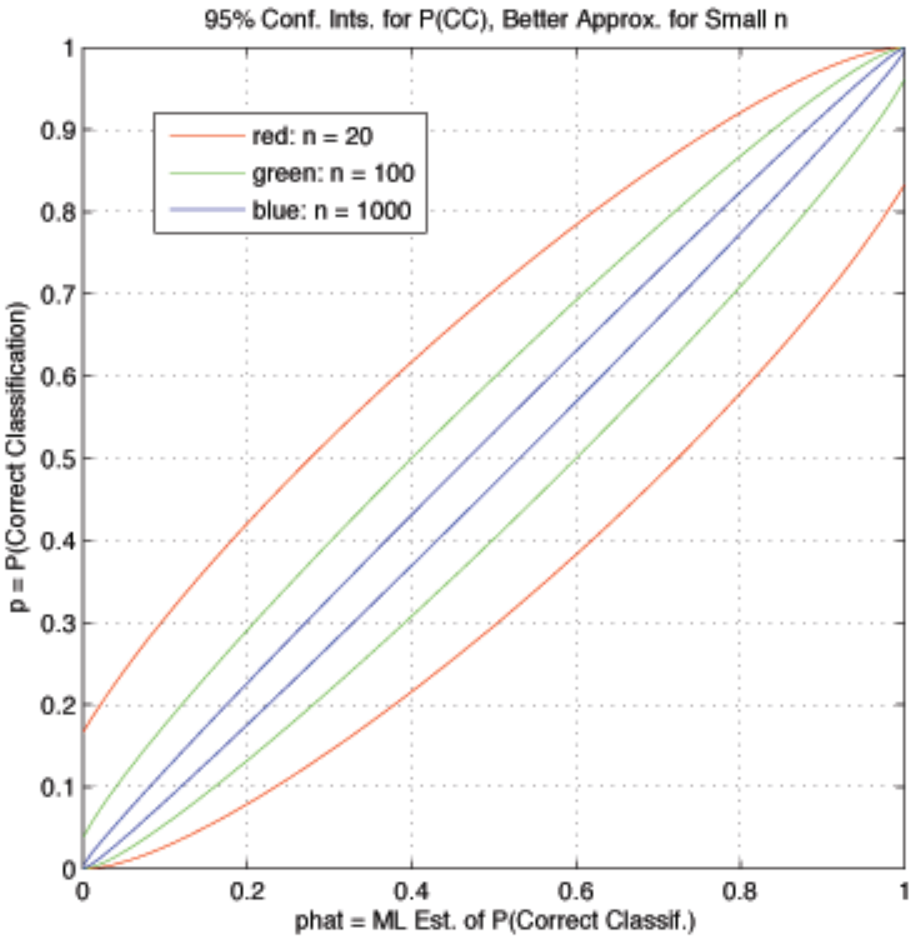


Figure 7.9: Small Sample Size Case: The 95 Percent confidence interval bounds (L, U) for the probability of correct classification are plotted, given the sample size n. Here, we use the better approximation for the the case in which n is small. Note how the confidence interval tightens as the sample size increases.

Chapter 8

MATLAB Software for Cable Damage Detection

A set of MATLAB software packages was created to implement and demonstrate the algorithms for this project. They are described in this section. They are also collected together in a set of folders and on a compact disc (CD). The core signal processing algorithms for the following codes were written by G. Clark as "research code "(see the next section). The executable (deliverable) codes, GUI's and high-level codes that package the signal processing codes and make them useable by others were written by K. Wade. The RepeatabilityTest.m code started as a code by Jessie A. Jackson. Then they were modified and extended by C. L. Robbins, with some input from G. A. Clark.

The user can obtain additional information about the codes by reading the headers for the files and the comments inside the code. For example, the headers describe the inputs, outputs and functions called.

8.1 Research Code Written by G. A. Clark

8.1.1 Signal Pre-Processing Algorithms

Cut_TDR_Ref_GAC_v2.m

Given an ensemble of TDR signals, this function computes ensemble averages, and allows the user to cut and zero-pad a reference signal $s(n)$ interactively.

Cut_TDR_Sig_GAC_v1.m

Given an ensemble of TDR signals, this function computes ensemble averages, and allows the user to cut and zero-pad a measurement signal $x_U(n)$ interactively.

Cut_TDR_Damage_GAC_v1.m

Given an ensemble of TDR signals, this function computes ensemble averages, and allows the user to cut and zero-pad an artificial damage wavelet $d(n)$ interactively.

Make_xd_GAC_v1.m

Given an ensemble of TDR signals, this function computes ensemble averages, and allows the user to create an artificial damage signal $x_D(n)$ interactively. This signal consists of a damage wavelet $d(n)$ added to a known undamaged signal $x_U(n)$.

Pre_MBD3_GAC_v2.m

Given the three signals $s(n)$, $x_U(n)$ and $x_D(n)$ created above, this code demeans, detrends, delays and decimates the signals according to parameters chosen by the user. This is done to prepare the signals for the System Identification step.

8.1.2 System Identification Step

The user runs the System Identification Toolbox Graphical User Interface (GUI) to build an ARX model for the cable system, using the $s(n)$ and $x_U(n)$ signals above. This step does not involve any codes written at LLNL. The model $ARX(n_a, n_b, n_k)$ is specified by a set of three parameters, $n_a =$ the order of the $A(q^{-1})$ polynomial, $n_b =$ the order of the $B(q^{-1})$ polynomial, and $n_k =$ the integer number of delay samples estimated between the input and output. These parameters are passed to the model-based detection codes below.

8.1.3 Model-Based Damage Detection Algorithms

MBD_GAC_v4.m

Given the ARX model $ARX(n_a, n_b, n_k)$ estimated above and the three pre-processed signals from above, this code computes the innovations and various other associated signals and correlations. It then allows the user to test the innovations for statistical whiteness using both the Zero Mean Whiteness testing code and the WSSR testing codes below. The final output is a decision as to whether or not the given $x_D(n)$ signal under test represents a damaged cable or not.

Test_White_GAC_v2.m

This function implements the zero-mean whiteness test; then it produces a plot and a decision about whether the innovations are white or not.

WSSR_Test_GAC_v2.m

This function implements the Weighted Sum Squared Residuals test for the innovations and prints out results in the form of a plot and a decision as to whether or not the WSSR test passes or fails.

8.1.4 Time Delay Estimation and Signal Alignment Algorithms

Delay_Est_GAC.m

Given a reference signal and a signal under test, this function estimates the delay (or advance) between them. This is done by finding the delay at which the largest peak in the cross-correlation function occurs.

Align_Signals_GAC.m

Given two signals and the estimated delay (or advance) between them, this function shifts the signal under test by the amount of the estimated delay.

Auto_Align_Signals_GAC.m

Given two signals, this function calls the two functions above to estimate the delay between the signals and align them.

Align_Signals_Caller_GAC.m

This is a high-level code designed to set up a problem and call the auto-align code for demonstration purposes.

8.1.5 Receiver Operating Characteristic Curve and Statistical Confidence Interval Algorithms

ROC_Driver_GAC_v1.m

Given an ensemble of declaration vectors d_U for the H_0 hypothesis and an ensemble of declaration vectors d_D for the H_1 hypothesis, this code computes the probability of detection and probability of false alarm for the detector/classifier.

Conf_Int_Binomial_GAC.m

Given N, the integer number of statistical samples in a statistical test, and the estimated probability of correct classification estimated for the test, this function computes the confidence interval bounds for an assumed significance of 5 percent. Two estimates of the bounds are computed. The first is under the assumption that N is greater than 30, so the underlying distribution can be assumed Gaussian. The second is under the assumption that $N < 30$ so the "exact "estimate is computed without the Gaussian assumption.

Conf_Int_Bin_Caller_GAC.m

This is a high-level code designed to set up a problem and call the confidence interval code above for demonstration purposes.

8.1.6 General Signal Processing Algorithms

Make_Time_Vector_N_GAC.m

This function creates a vector of time values that can be used for plotting a signal with the proper time scaling.

power_of_two_GAC.m

Given a positive integer, this function computes the next highest power of two. This is generally used when computing a fast fourier transform and the user needs to find an FFT size (a power of two) that is larger than the signal length.

8.2 Deliverable Software Written by K. A. Wade

This set of software consists of executable code deliverable to the sponsor.

LAUNCH FILETESTER GUI
Launches filetester_gui in the /Filetester_gui directory.
This is an interactive gui for tweaking the parameters by which a file is tested.

LAUNCH FILETESTER EXE
Launches the executeable version of the filetester as it would be done in DOS at the command line.
This approximates what Pantex will receive and use.
(Note: the Filetester_exe directory also contains the compiled code and resulting artifacts and subfolders).

8.3 Software for General Use Written by K. A. Wade, G. A. Clark and C. L. Robbins

This set of software is designed for general use. The organization of the codes is depicted in Figures (8.1),(8.2),(8.3), and (8.4). A stripped signal file is one with the header removed. tdr_menu is a top level ”menu” code for running all the codes described below. The word “LAUNCH”in the code descriptions below refers only to the labels on the GUI buttons for the various codes. “LAUNCH”is just an instruction to the user and is not a part of the file name(s).

CODES

LAUNCH ROC ENSEMBLE BUILDER PROGRAM
Launches ROC_Ensemble_Builder.m in the /ROC_code directory.
This is used for creating ROC curves for a particular data set.

DOCUMENTATION

README:
Launches this file (and maybe others in future)

DATA CHECKING

LAUNCH REPEATABILITY GUI
Launches RepeatablityTest.m in the /RepeatabilityAnalysisCodesV1 directory.
This code is by Chris Robbins et. al.

LAUNCH SIGNAL CONDITIONING TOOLS MENU

The tools under this GUI button are summarized in Figure (8.3).

AVERAGE
Launches average_gui in the /Signal_Manipulation directory.
Allows user to take the average of several original data files.

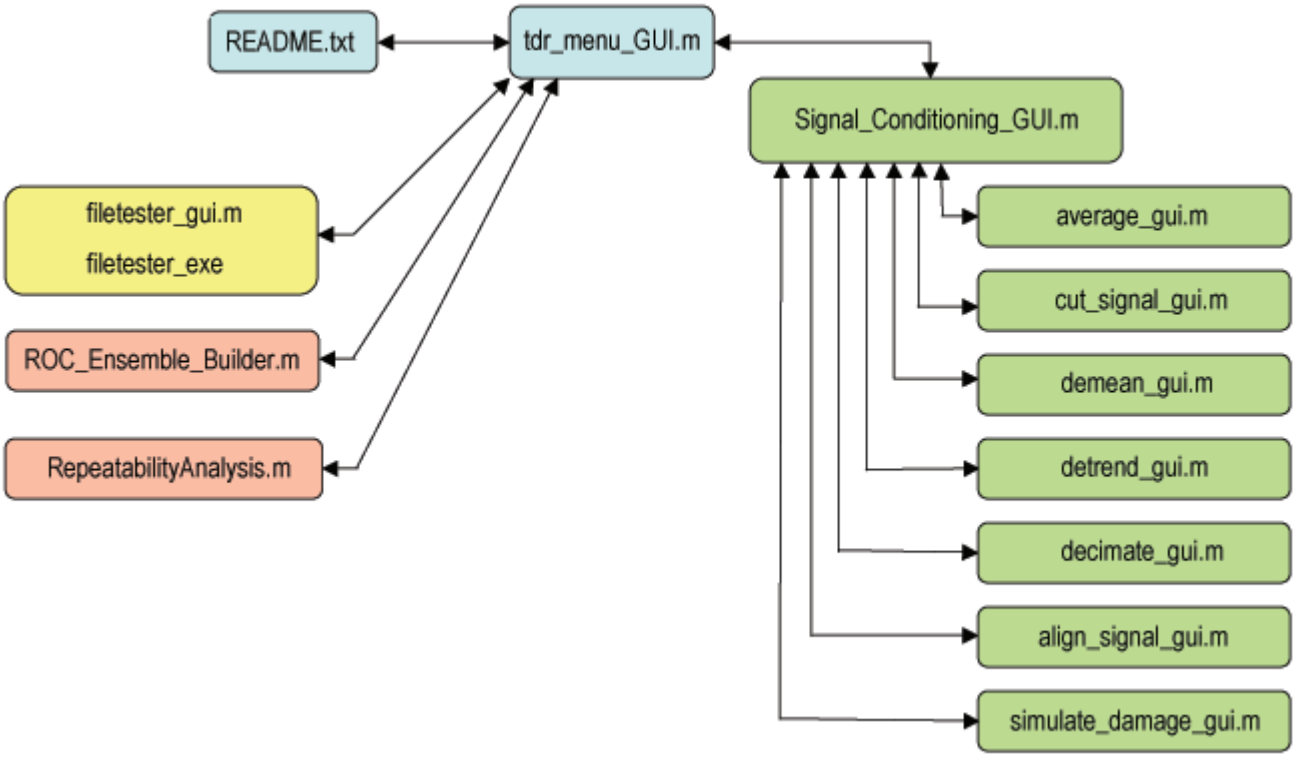


Figure 8.1: Deliverable Software and TDR Software for General Use: A set of MATLAB tools, including GUI's has been created to implement the algorithms needed for cable damage detection. These codes are deliverable to the sponsor. This figure depicts the codes with their names and their interactions.

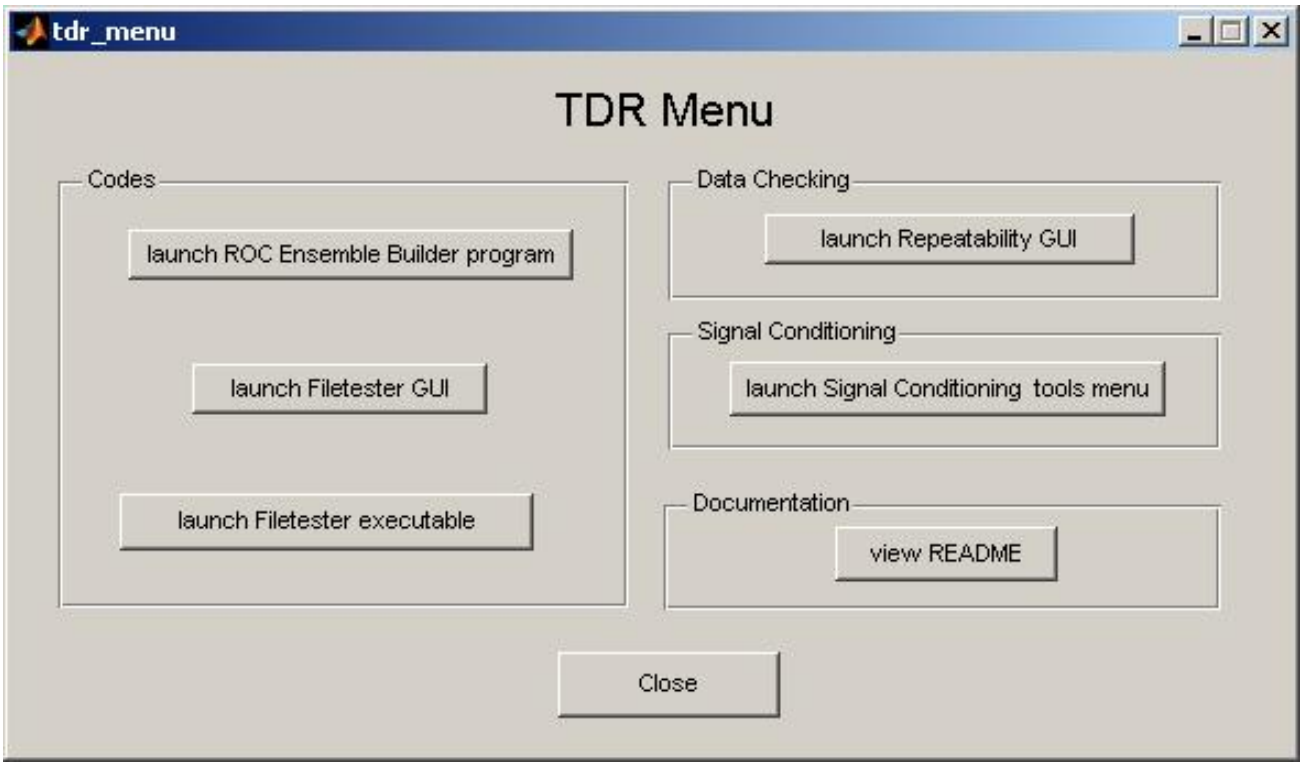


Figure 8.2: tdr menu screenshot: This figure depicts a representative image of the tdr menu GUI screen. Note that the buttons launch the codes depicted in Figure (8.1).
.

Saves the average file without its header for use with the other tools.

CUT
Launches cut_signal_gui.m in the /Signal_Manipulation directory.
Allows user to cut a stripped signal.

DEMEAN
Launches demean_gui.m in the /Signal_Manipulation directory. Allows user to demean a stripped signal.

DETREND
Launches detrend_gui.m in the /Signal_Manipulation directory. Allows user to demean a stripped signal.

DECIMATE
Launches decimate_gui.m in the /Signal_Manipulation directory.
Allows user to decimate or interpolate a stripped signal by an integer or fractional value.

ALIGN
Launches align_signal_gui.m in the /Signal_Manipulation directory.
Allows user to align two stripped signals. See Figure (8.4).

SIMULATE DAMAGE
Launches simulate_damage_gui.m in the /Signal_Manipulation directory.
Allows user to create an artificial ”damaged” signal by cutting and superimposing one stripped signal over another stripped signal.

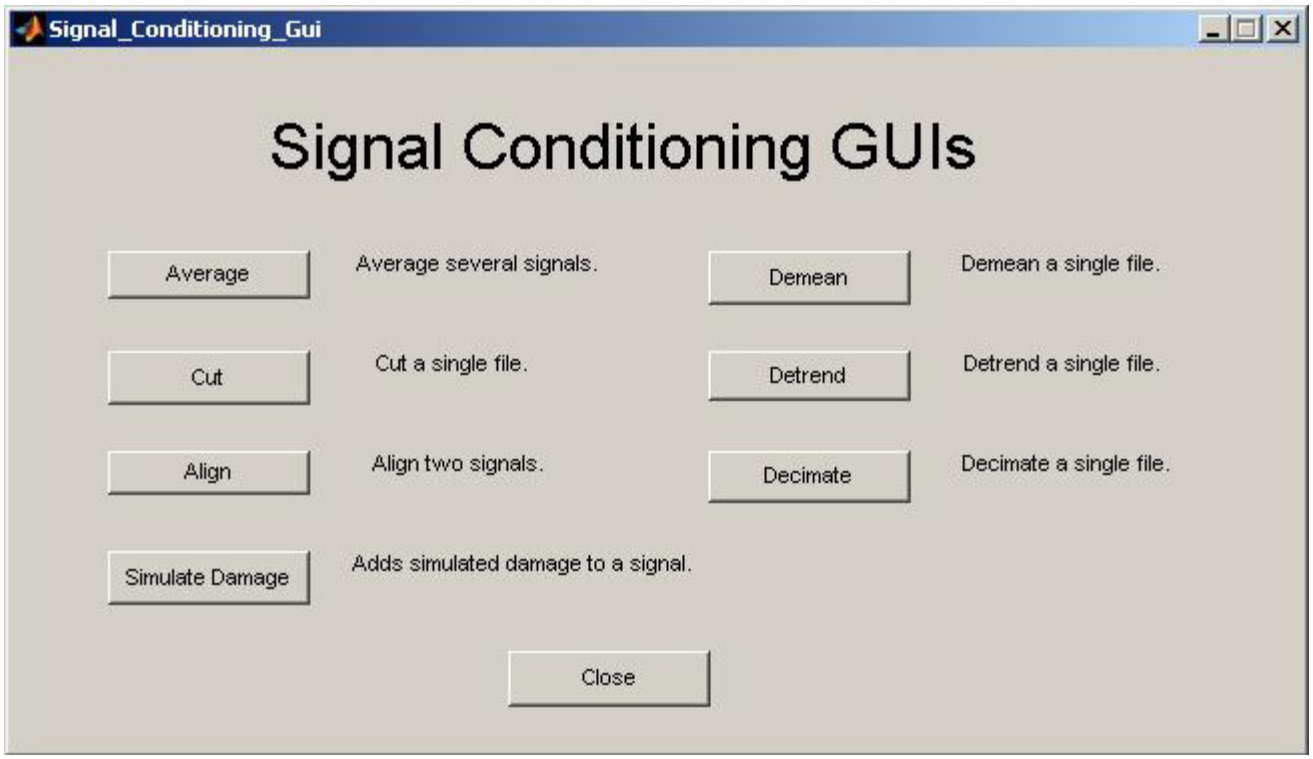


Figure 8.3: Signal Conditioning Screenshot: This figure depicts a representative image of the signal conditioning GUI screen. Note that the buttons launch the codes depicted in Figure (8.1). Activate this GUI by selecting the button marked LAUNCH SIGNAL CONDITIONING TOOLS MENU.

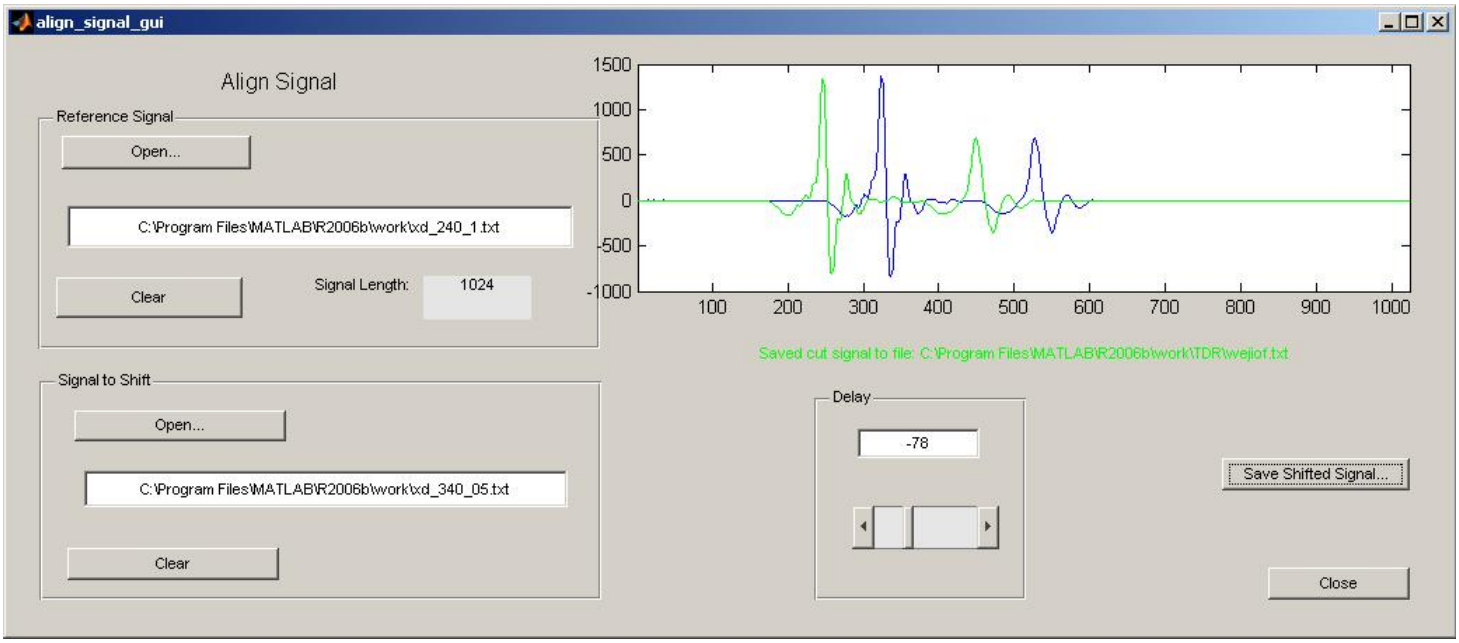


Figure 8.4: Align GUI Screenshot: This figure depicts a representative image of the align GUI screen. Note that the buttons depicted in Figures (8.2) and (8.3) launch the codes depicted in Figure (8.1). The other signal conditioning GUI's listed in Figure (8.1) are similar to this, but not depicted in the report to save space.

Chapter 9

Experiment E1: Preliminary Experimental Results with an Ensemble of 12 Signals

This section represents preliminary results, in the sense that it used a preliminary data set acquired to assess the experimental procedures, hardware, software, etc. It does not represent the final results. At the time the data were collected, the PIU was in the process of being modified to improve the attachment repeatability. Therefore, the PIU was not removed between data acquisitions in order to obviate any repeatability effects associated with the PIU attachment. The types of damage are summarized in Figure (9.1). This experiment occurred before the 2D test fixture was available, so the cables were mounted flat on a lab bench during testing.

Three signals were collected for each damage type. The names of the signal files are shown in the figure (minor1a, etc.). There were three main damage categories: Undamaged, Minor Damage and Major Damage. For the Undamaged and Major Damage cases, there were no damage subsets. For the Minor Damage case, there were three subsets (Minor 1, Minor 2 and Minor 3) as shown in Figure (9.1).

In hindsight, it might have been better to have named the damage categories differently. The terms minor and major were actually intended to describe the manifestation of abnormalities in the TDR signal, not the damage to the cable itself. The key result of the test is that if the conductors are shorted, the damage manifestation in the TDR signal is obvious. If the conductors are not shorted, the damage manifestation in the TDR signal is not obvious. However, because these results under the original damage names have been committed to documentation in several papers, reports and software texts, we choose to continue to use them in this report. The reader can make any adjustment desired.

Note that the ROC Curves are not smooth because we had only a few sample points with which to estimate the P_D and P_{FA}. The small sample size comes from the fact that we had available only a few undamaged cable signals and a few damaged cable signals. The signal pre-processing steps for the System Identification Step are summarized in Figure (9.2). This is a sub-optimal situation. Because we did not have much data, so we were forced to do something we do not like to do: Use the same signal to create both $s(n)$ and $x_U(n)$. The signal pre-processing steps for the cable testing step are shown in Figure (9.3). Figure (9.4) depicts three of the basic signals used in the analysis. The top two signals are the $s(n)$ and $x_U(n)$ signals used to build the model in the System Identification step. The bottom signal is plotted to show a representative damage signal. It is the signal for the Major Damage case.

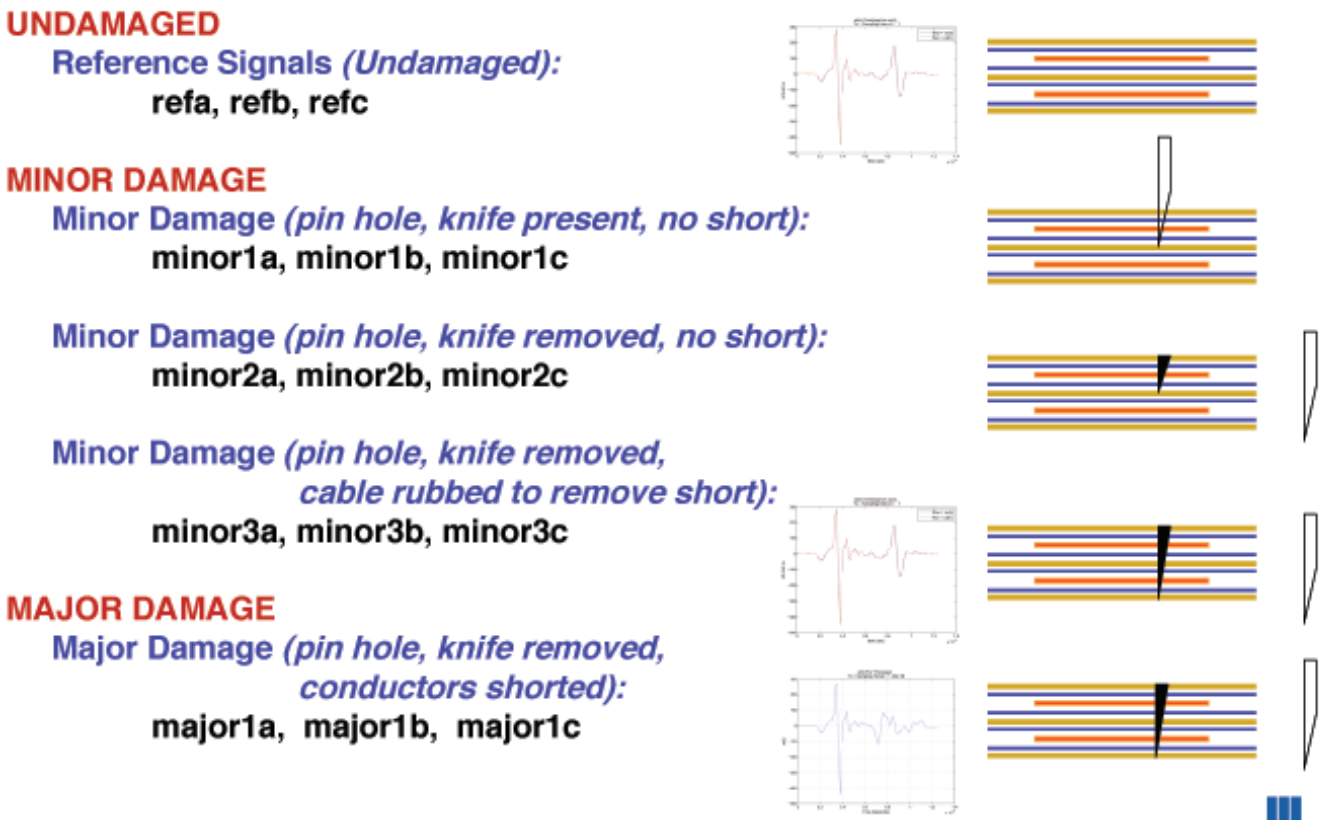


Figure 9.1: E1: Twelve signals were acquired using a set of experiments to create artificial damage into the cables.

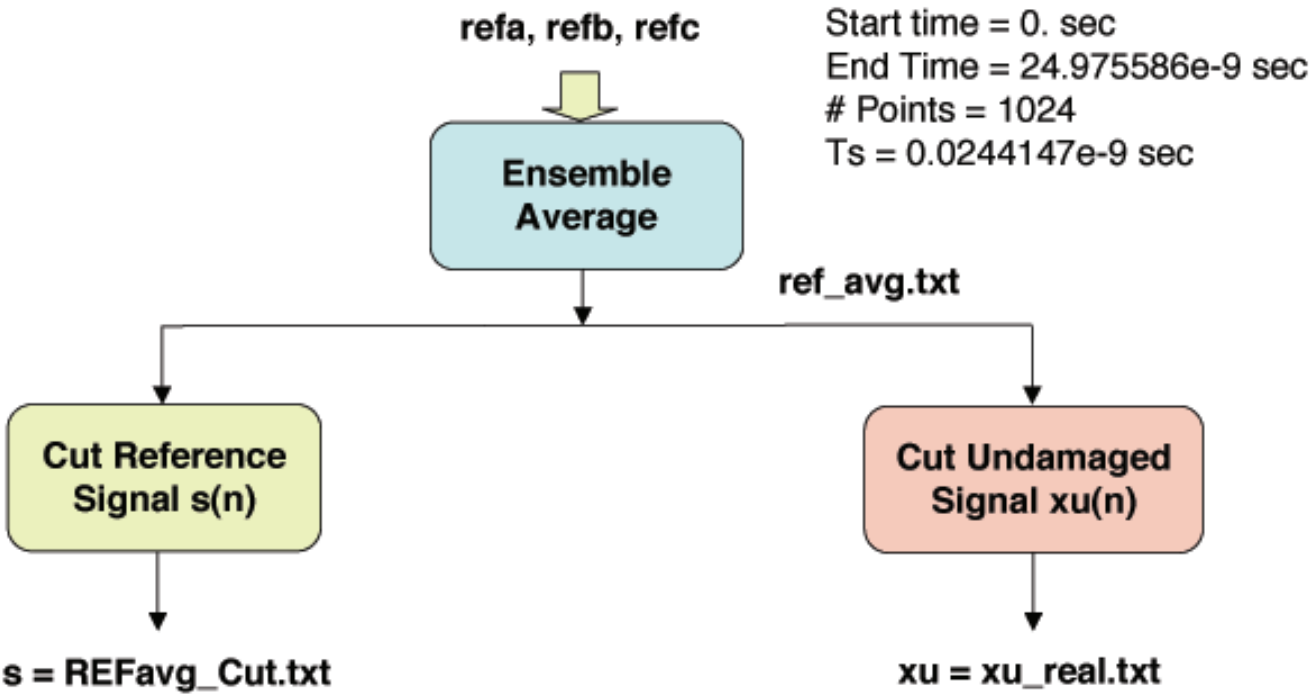


Figure 9.2: E1: Undamaged signals were cut for Step1, System Identification

9.1 The System Identification Step

A model was identified and is specified by $arx(n_a, n_b, n_k) = arx(18, 22, 21)$. Figure (9.5) shows that the autocorrelation of the innovations lies within the confidence bounds and the cross correlation R_{se} is approximately zero. We conclude that the model is valid. Figure (9.6) depicts the zero-mean whiteness test on the innovations for the modeling step. Clearly, the innovations pass the whiteness test, so the model is declared valid. For the system identification step, we see in Figure (9.7) that the WSSR test showed that the innovations are white, further validating our declaration that the model is valid.

9.2 Major Damage Case

For the major damage case, Figure (9.8) shows the signals $x_U(n)$ and $x_D(n)$ overlaid. The damage is obvious in this case. Figure (9.9) shows that the major damage innovations $e_D(n)$ are large and correlated. Figure (9.10) shows the WSSR test result for the major damage case. The WSSR indicates a model mismatch, so the cable is correctly declared to be damaged. The Receiver Operating Characteristic Curve (ROC) yields ideal performance. However, we note that because the statistical sample size is small, the confidence interval on the probability of correct classification is wide.

9.3 Minor3 Damage Case

The plot in Figure (9.12) shows that the minor3 damage is very difficult to distinguish from visual inspection of the plot of $x_U(n)$ overlaid on $x_D(n)$. The plot in Figure (9.13) shows that the innovations are small, but correlated.

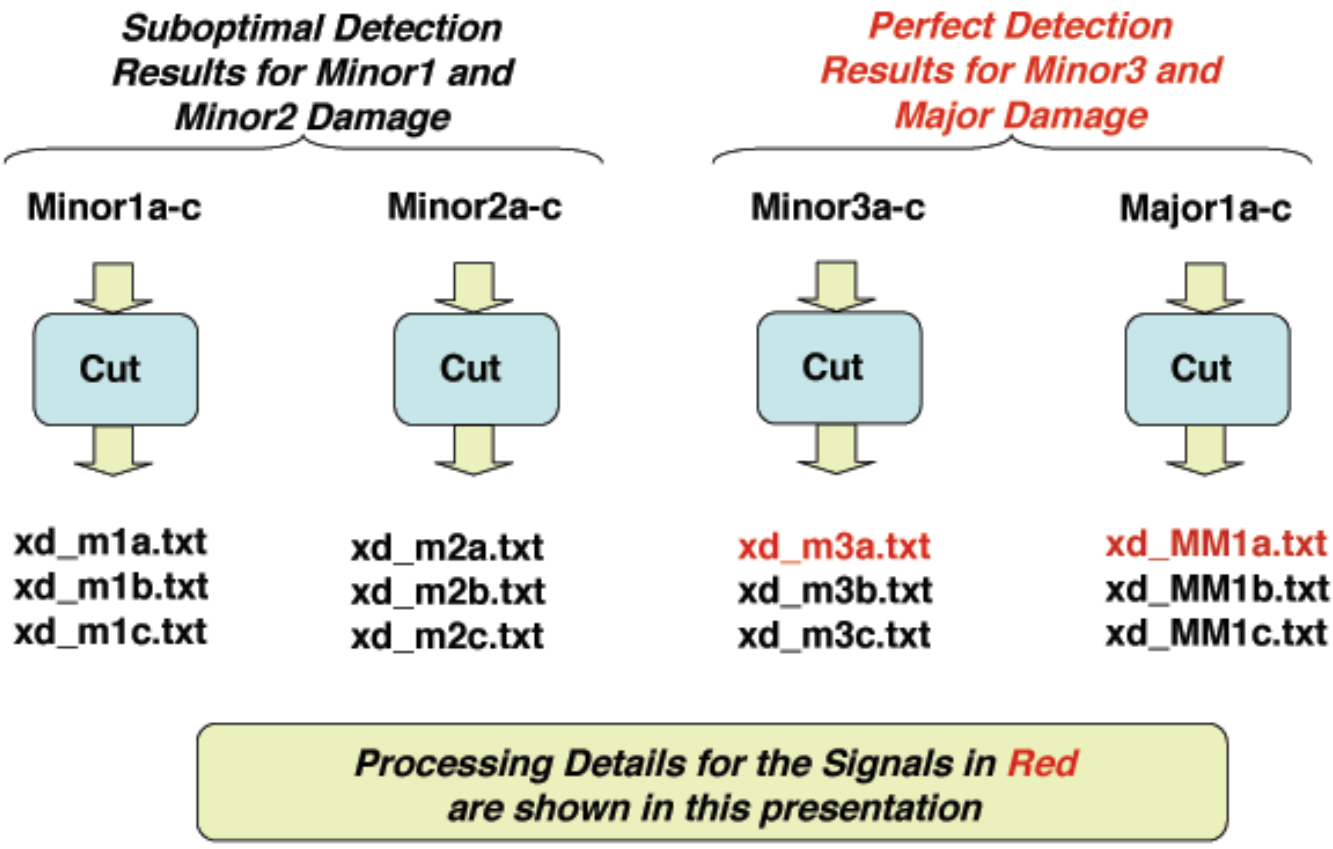


Figure 9.3: E1: The damage signals were cut for Step2, damage testing.

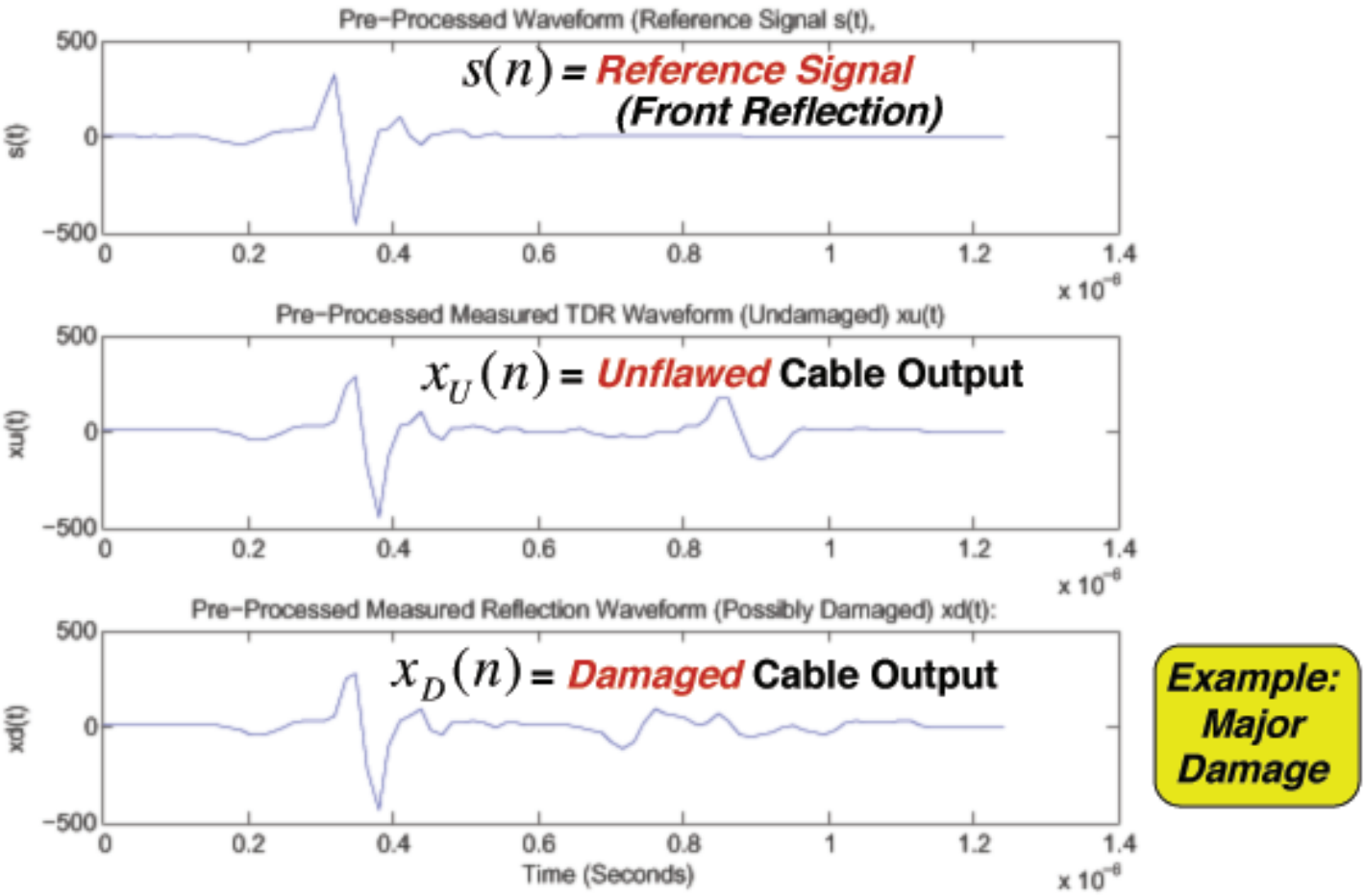


Figure 9.4: E1 Pre-Processed Signals: The top two signals are the $s(n)$ and $x_U(n)$ signals used to build the model in the System Identification step. The bottom signal is plotted to show a representative damage signal. It is the signal for the Major Damage case.

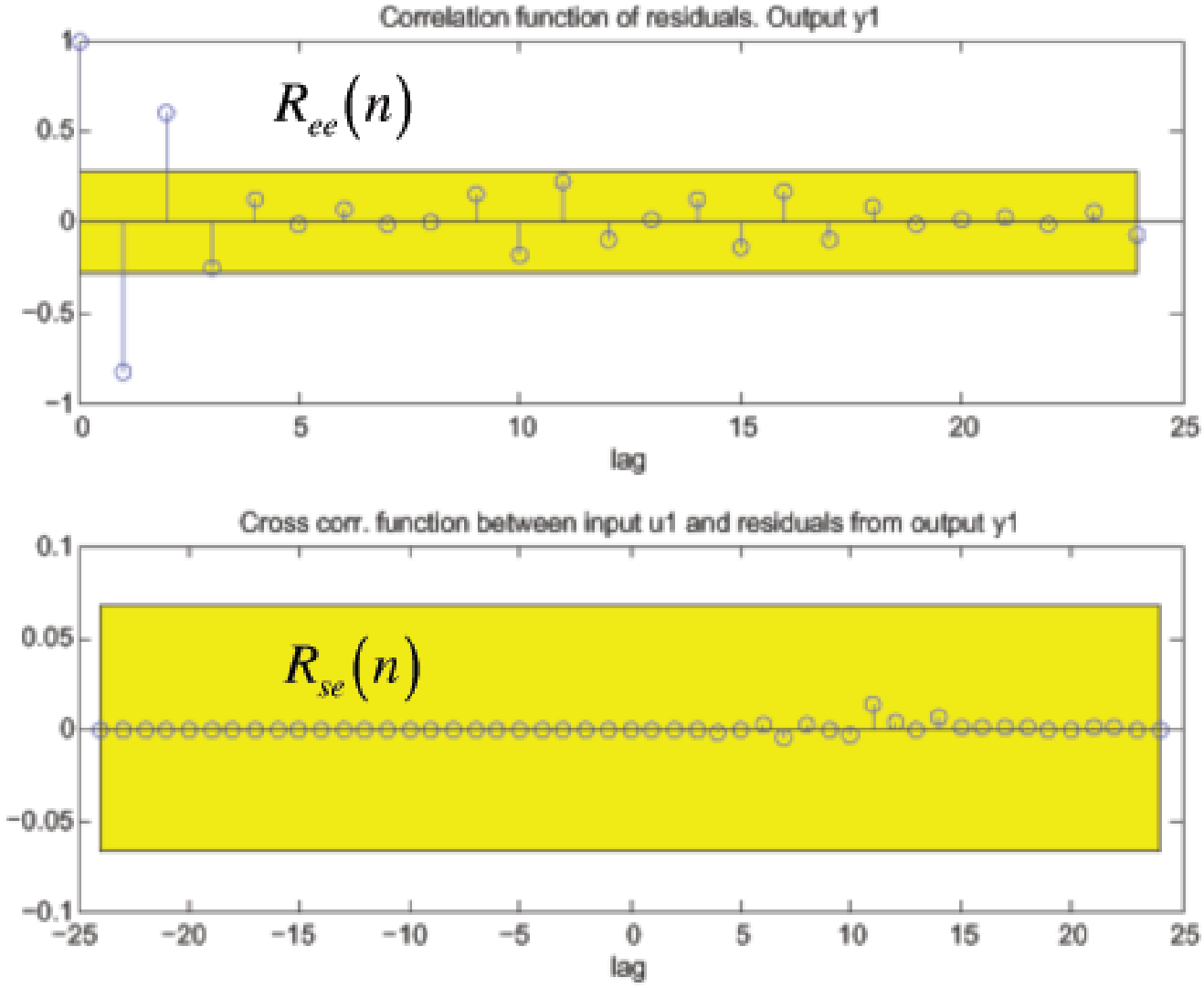


Figure 9.5: E1 System Identification Step: Correlation tests are satisfactory. The autocorrelation of the innovations (top) lie within the confidence bounds and the cross correlation R_{se} (bottom) is approximately zero. We conclude that the model is valid.

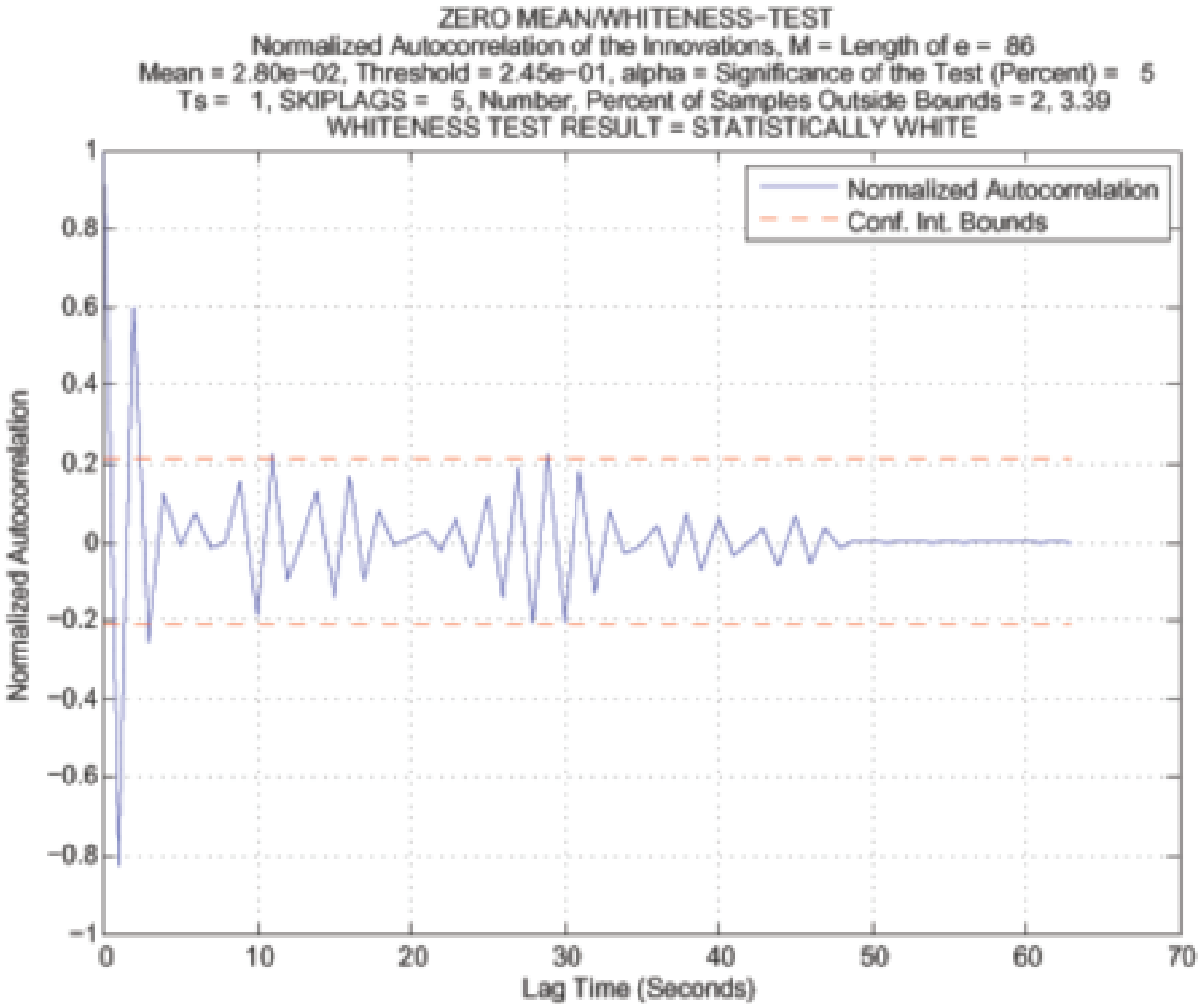


Figure 9.6: E1 System Identification Step: The Zero-Mean Whiteness Test on the model innovations shows that the innovations are white, and the model is valid.

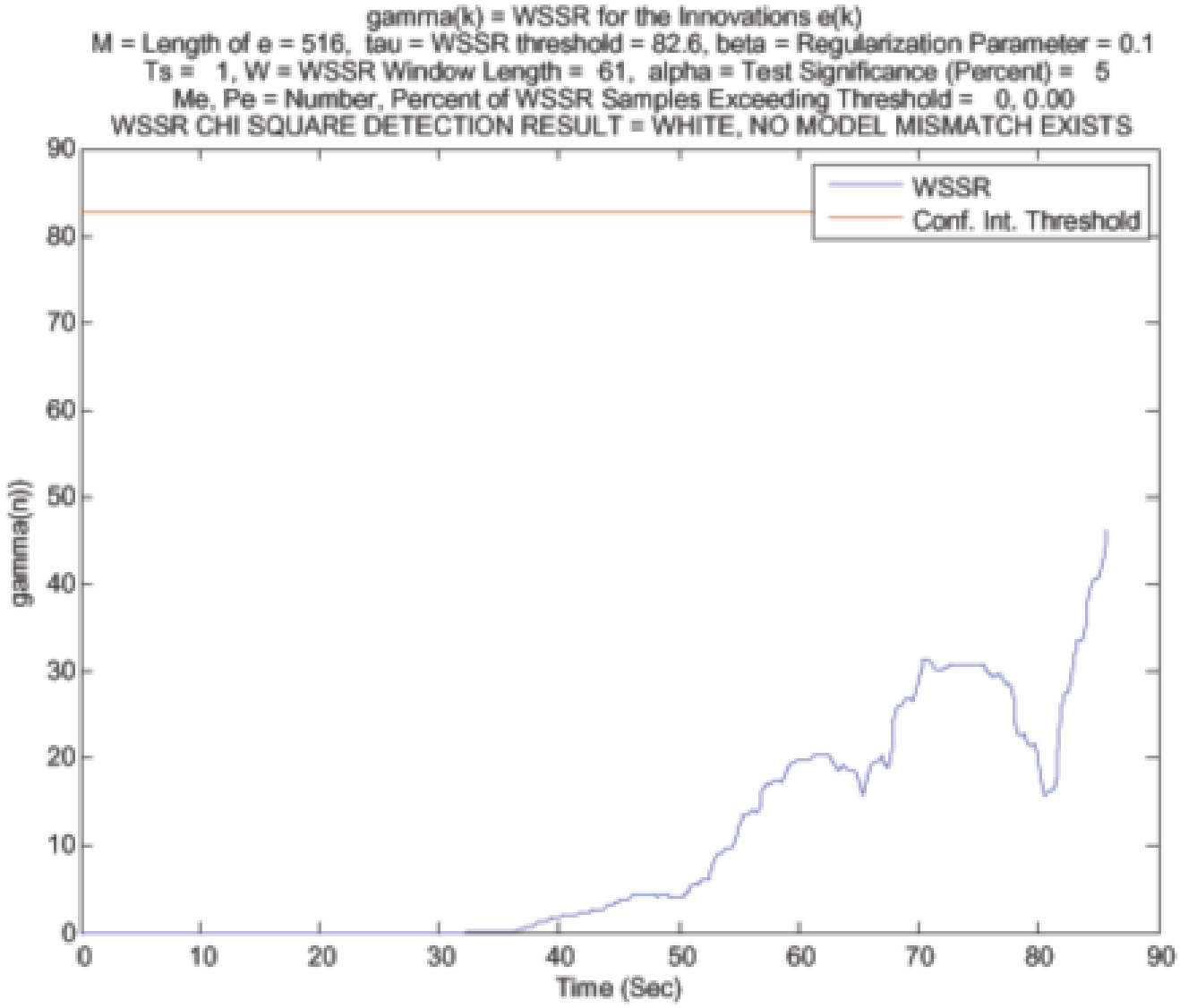


Figure 9.7: E1 System Identification Step: For the system identification step, we see that the WSSR test showed that the innovations are white, further validating our declaration that the model is valid.

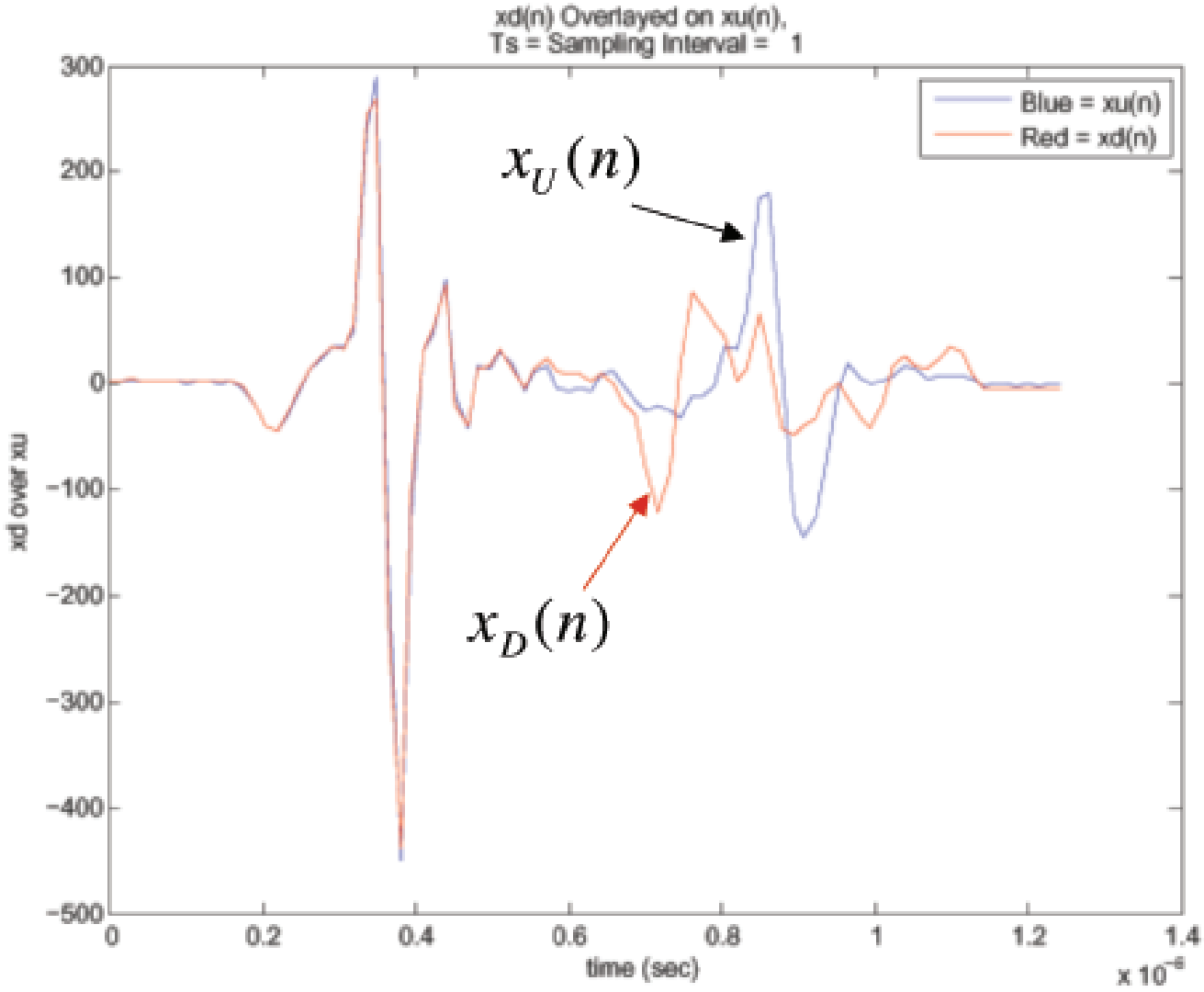


Figure 9.8: E1 Major Damage Case: The signals $x_U(n)$ and $x_D(n)$ are overlaid. The damage is obvious in this case.

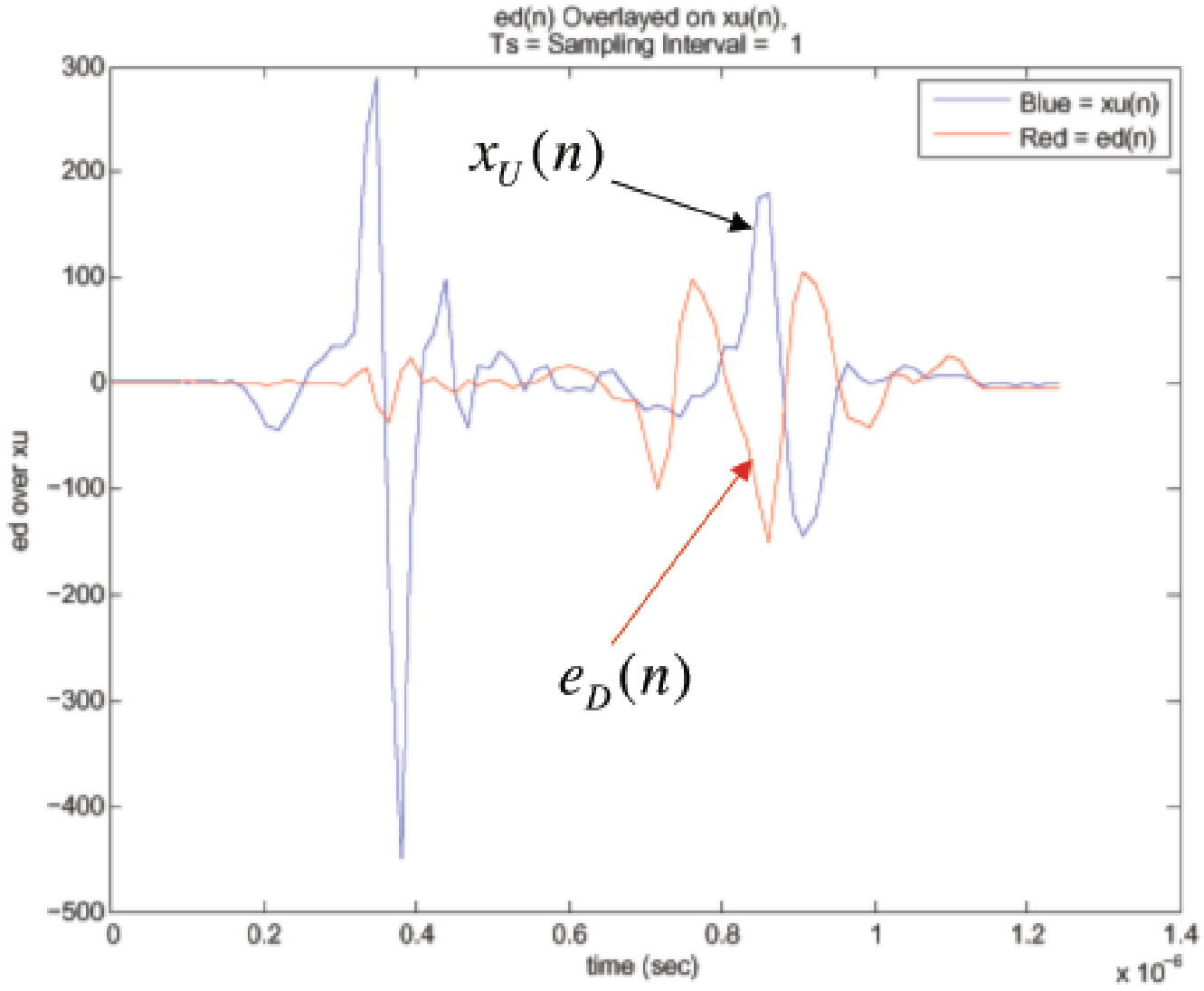


Figure 9.9: E1 Major Damage Case: Major damage innovations $e_D(n)$ are large and correlated.

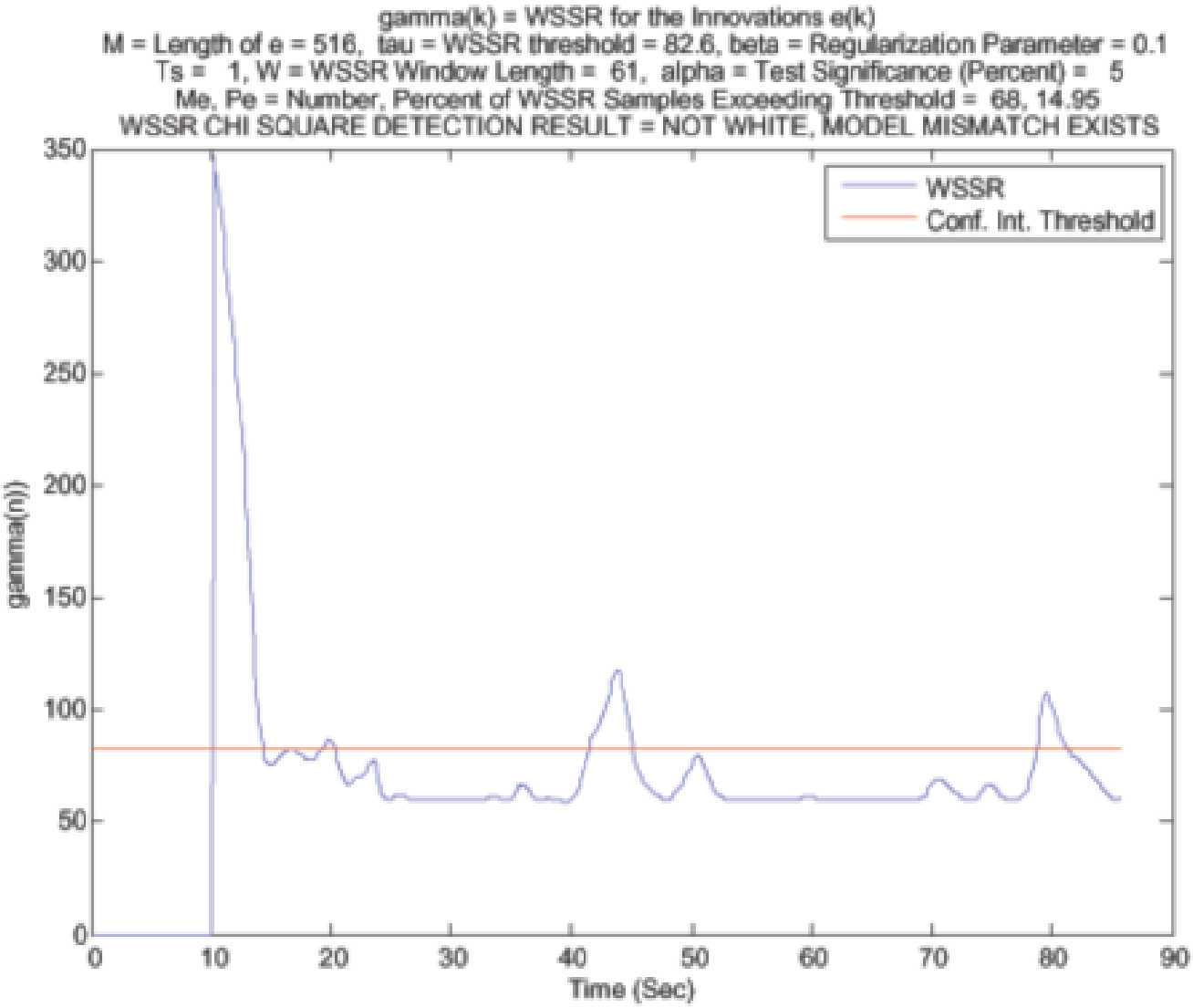


Figure 9.10: E1 Major Damage Case: The WSSR test indicates a model mismatch, so the cable is correctly declared to be damaged.

W	P_{FA}	P_D
24	1	1
36	0.66667	1
48	0.66667	1
49	0.66667	1
50	0.66667	1
51	0.66667	1
52	0.66667	1
53	0.66667	1
54	0.66667	1
55	0.66667	1
56	0.33333	1
57	0.33333	1
58	0.33333	1
59	0.33333	1
60	0	1
61	0	1
62	0	1
70	0	1
80	0	1

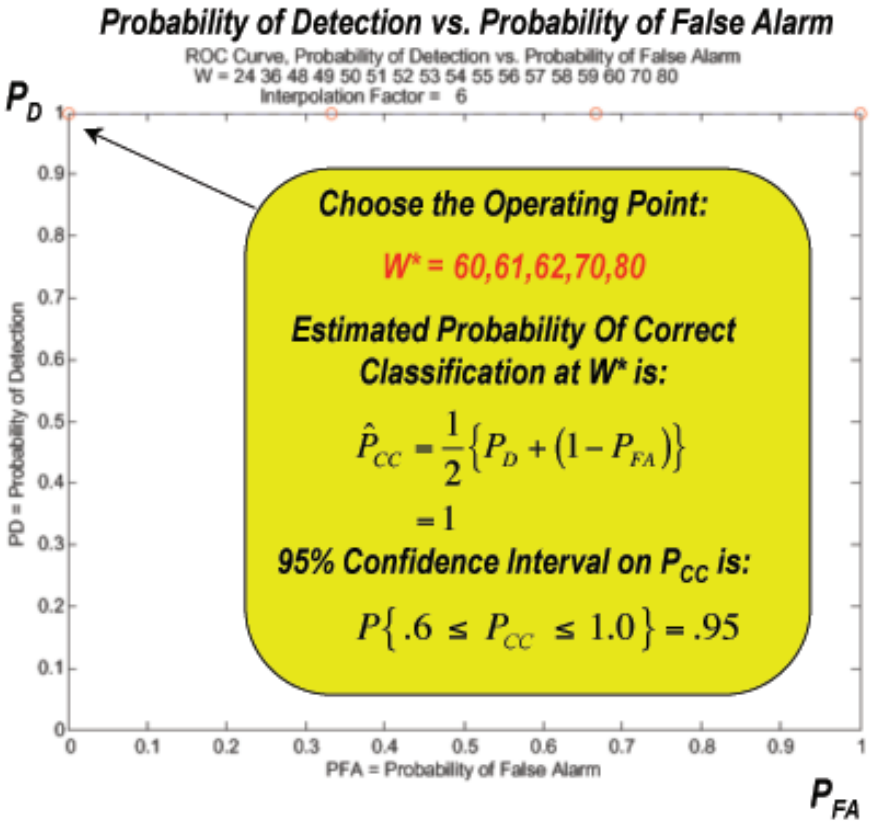


Figure 9.11: E1 Major Damage Case: The Receiver Operating Characteristic Curve (ROC) yields ideal performance. However, we note that because the statistical sample size is small, the confidence interval on the probability of correct classification is wide.

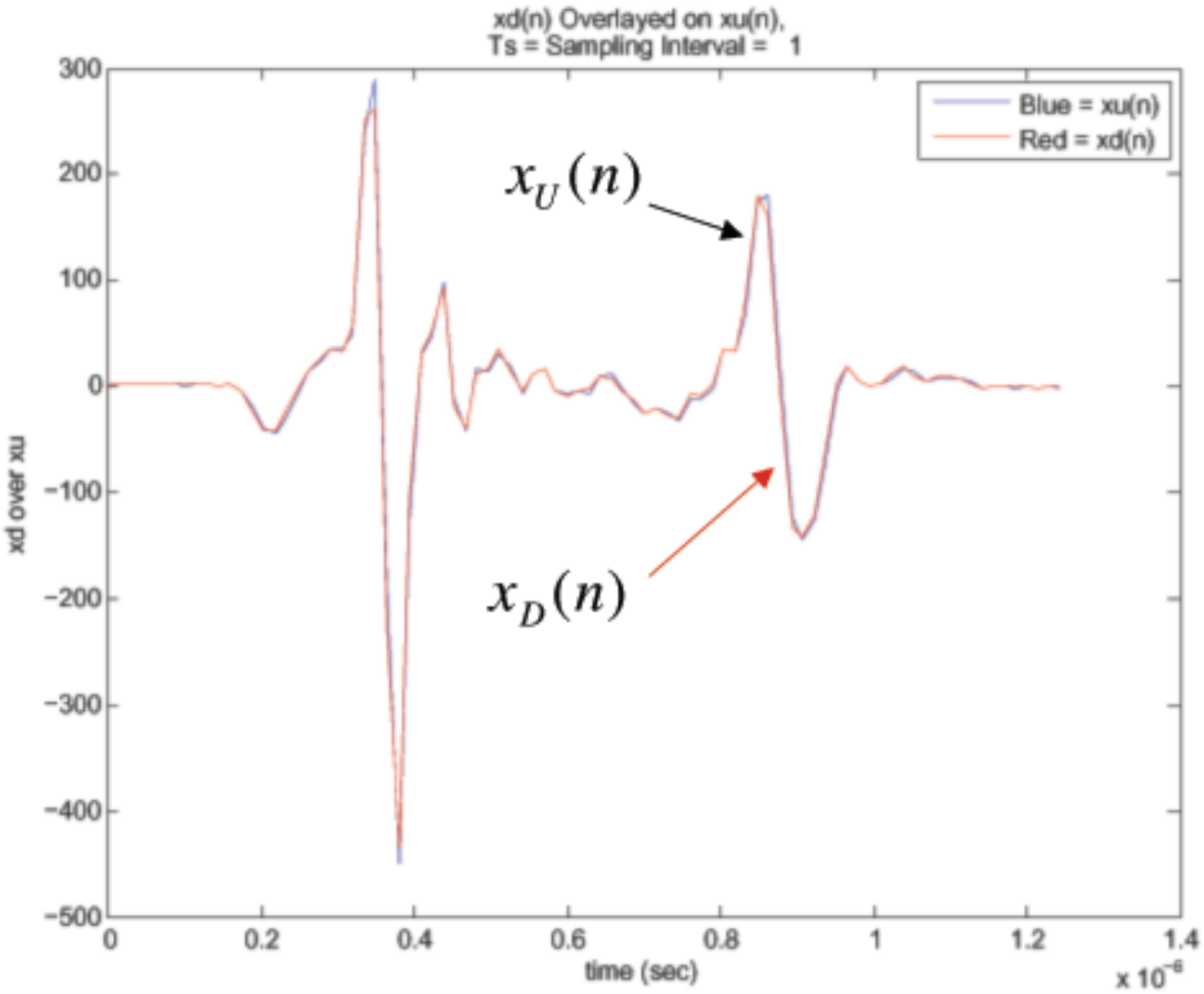


Figure 9.12: E1 Minor 3 Damage Case: The damage is very difficult to distinguish visually, as the signals $x_U(n)$ and $x_D(n)$ look almost identical.

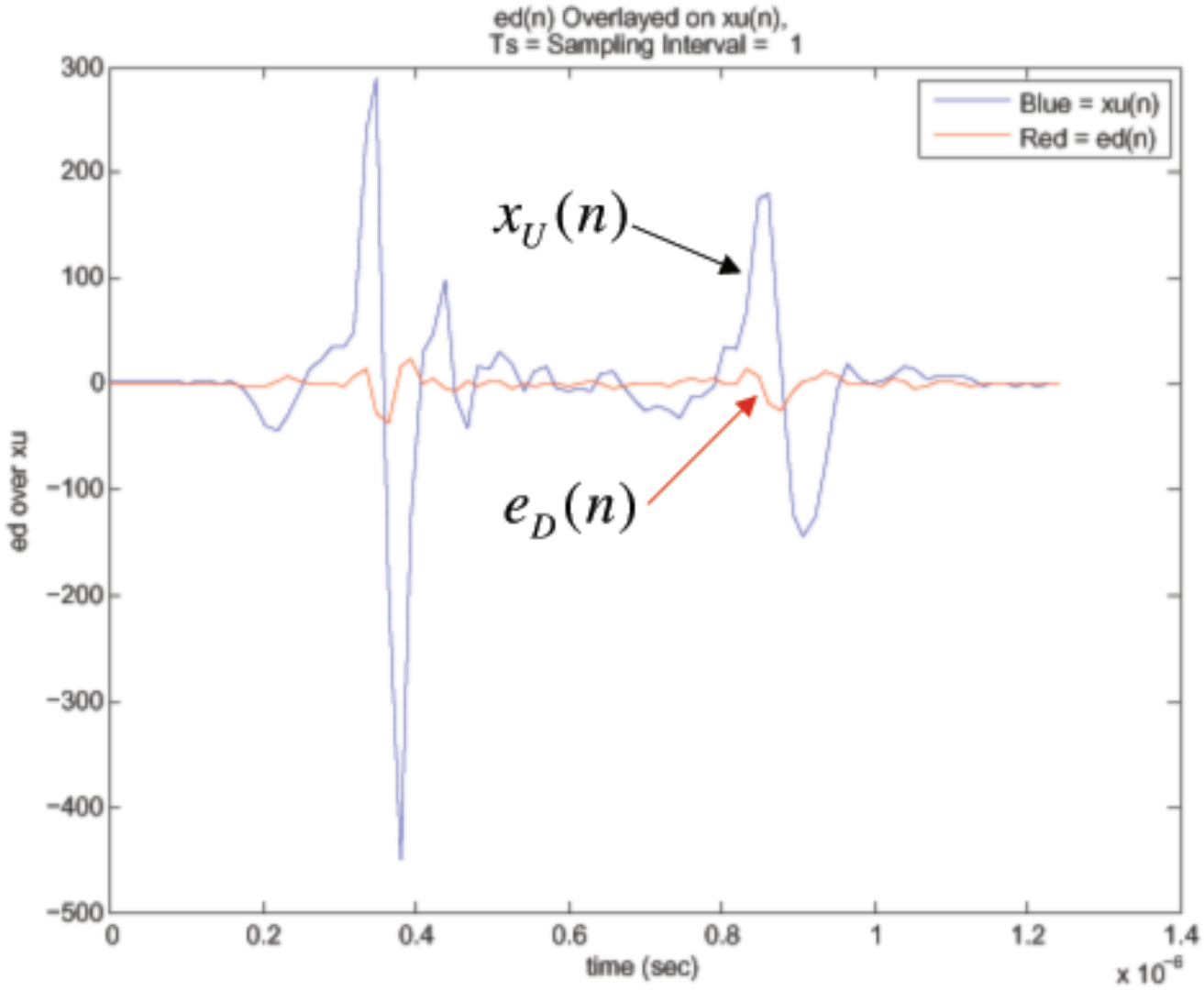


Figure 9.13: E1 Minor3 Damage Case: The innovations are small, but correlated.

Figure (9.14) shows that the WSSR test indicates a model mismatch, so the cable is correctly declared to be damaged.

9.4 Minor1a,b,c Damage Case

Figure (9.16) shows the ROC curve for the Minor 1 case. The Receiver Operating Characteristic Curve (ROC) yields an operating point with $P(CC) = .833$. We note that because the statistical sample size is small, the confidence interval on the probability of correct classification is wide.

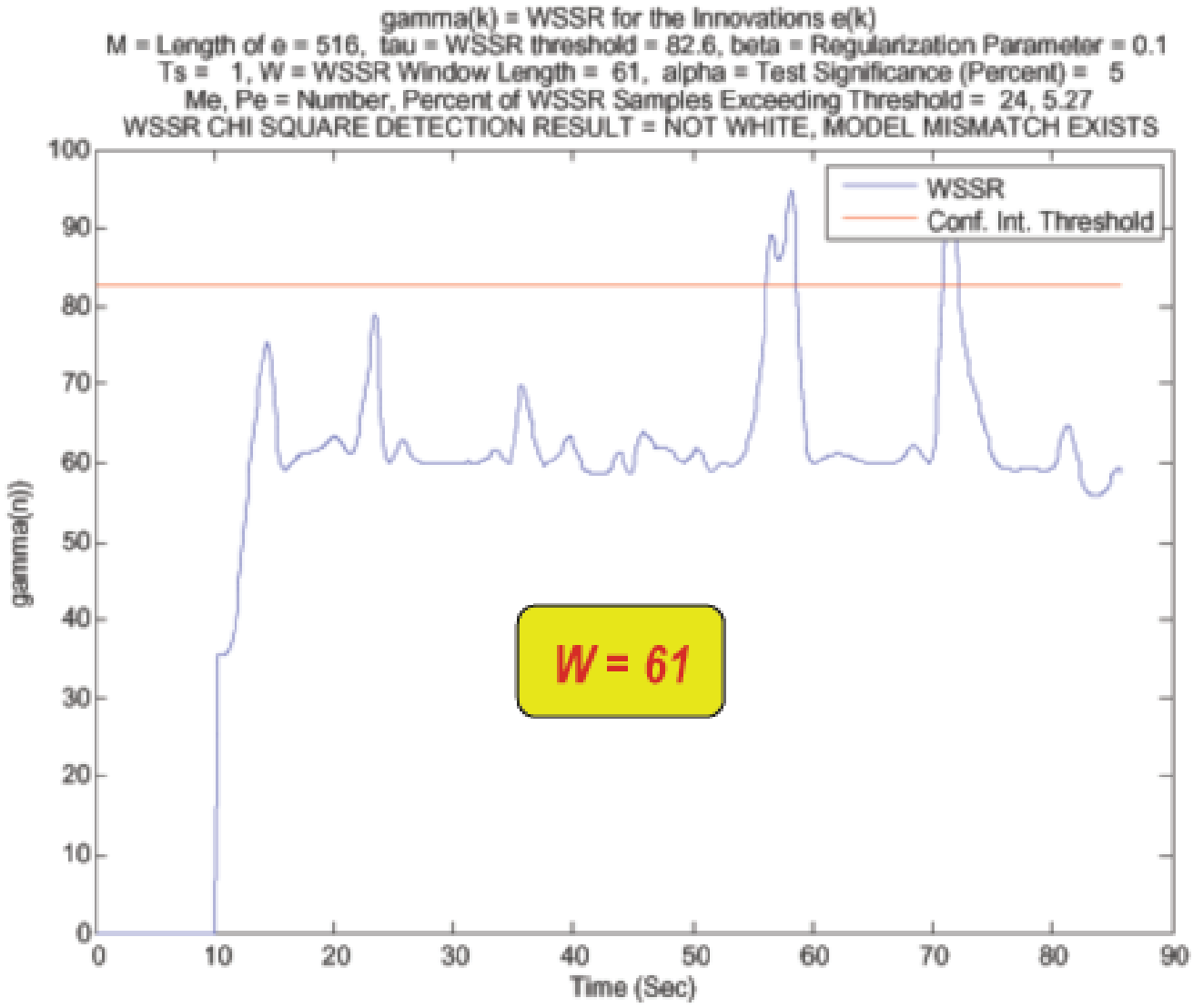


Figure 9.14: E1 Minor3 Damage Case: The WSSR test indicates a model mismatch, so the cable is correctly declared to be damaged.

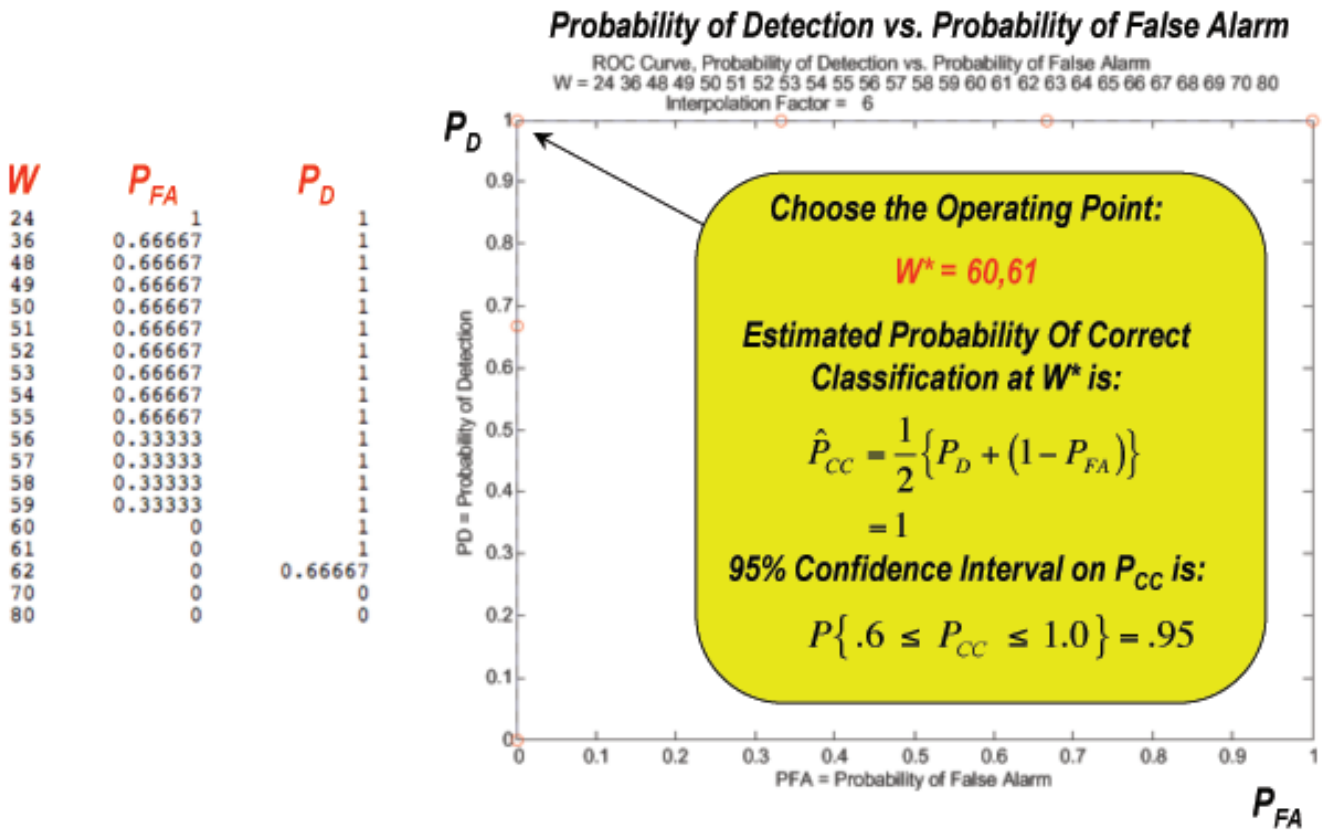


Figure 9.15: E1 Minor3 Damage Case: The Receiver Operating Characteristic Curve (ROC) yields ideal performance. However, we note that because the statistical sample size is small, the confidence interval on the probability of correct classification is wide.

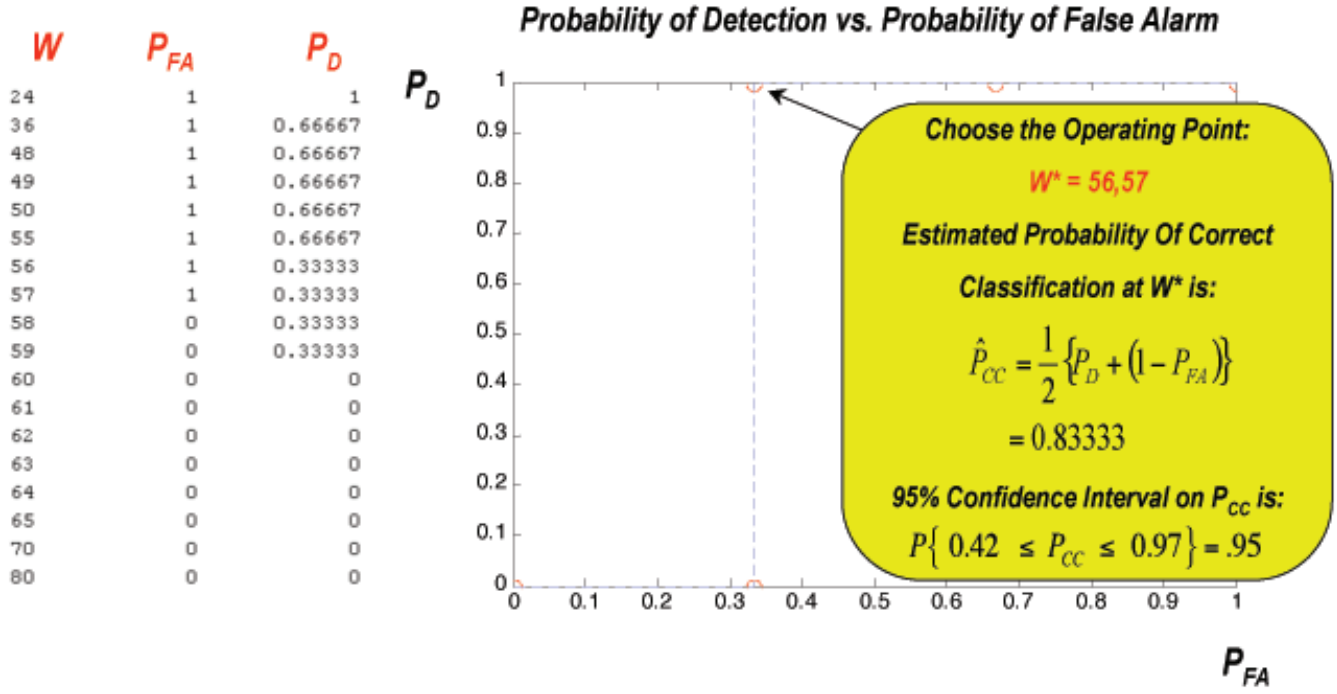


Figure 9.16: E1 Minor1 Damage Case: The Receiver Operating Characteristic Curve (ROC) yields an operating point with $P(CC) = .833$. We note that because the statistical sample size is small, the confidence interval on the probability of correct classification is wide.

9.5 Minor2a,b,c Damage Case

Figure (9.17) shows the ROC curve for the Minor 1 case. The Receiver Operating Characteristic Curve (ROC) yields an operating point with $P(CC) = .667$. We note that because the statistical sample size is small, the confidence interval on the probability of correct classification is wide.

9.6 All Damage Types Case

In this case, we were interested in seeing the detection performance if all of the damage types were lumped into one category and tested as such. Figure (9.18) shows the ROC curve for the All Damage Types case. The Receiver Operating Characteristic Curve (ROC) yields an operating point with $P(CC) = .792$. We note that because the statistical sample size is small, the confidence interval on the probability of correct classification is wide.

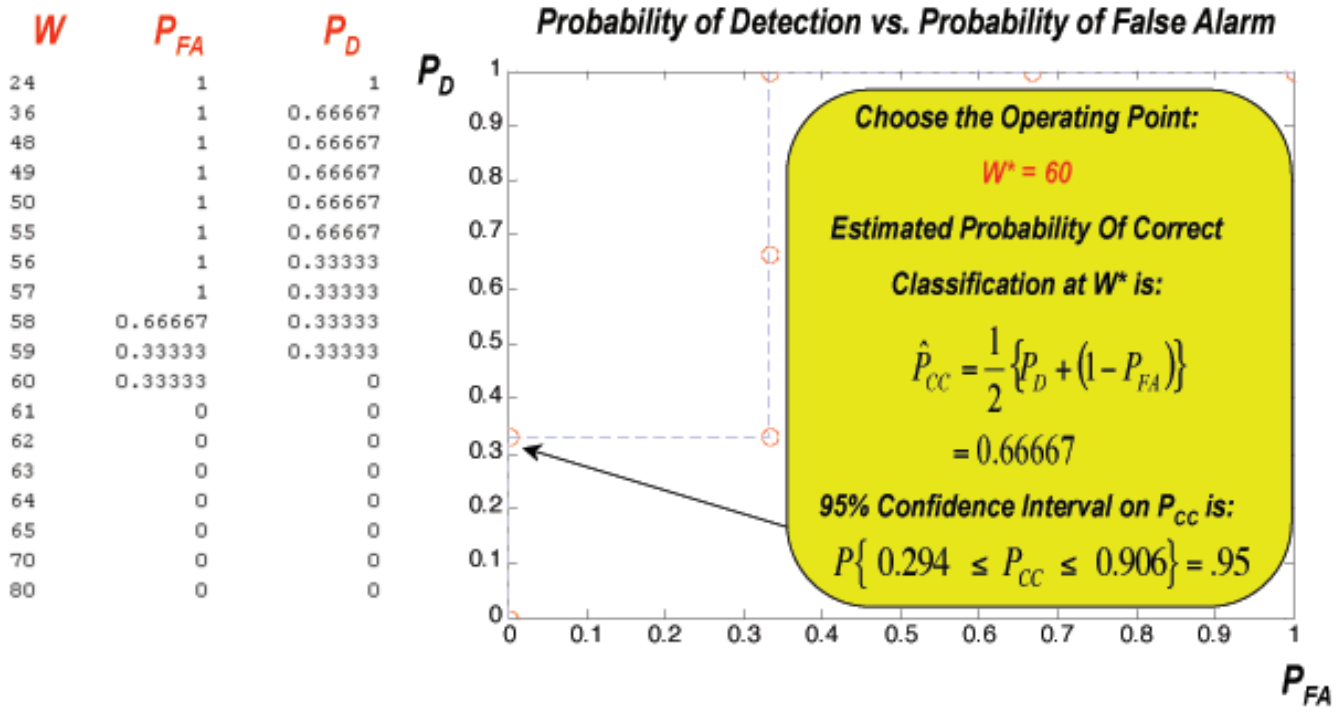


Figure 9.17: E1 Minor2 Damage Case: The Receiver Operating Characteristic Curve (ROC) yields an operating point with $P(CC) = .667$. We note that because the statistical sample size is small, the confidence interval on the probability of correct classification is wide.

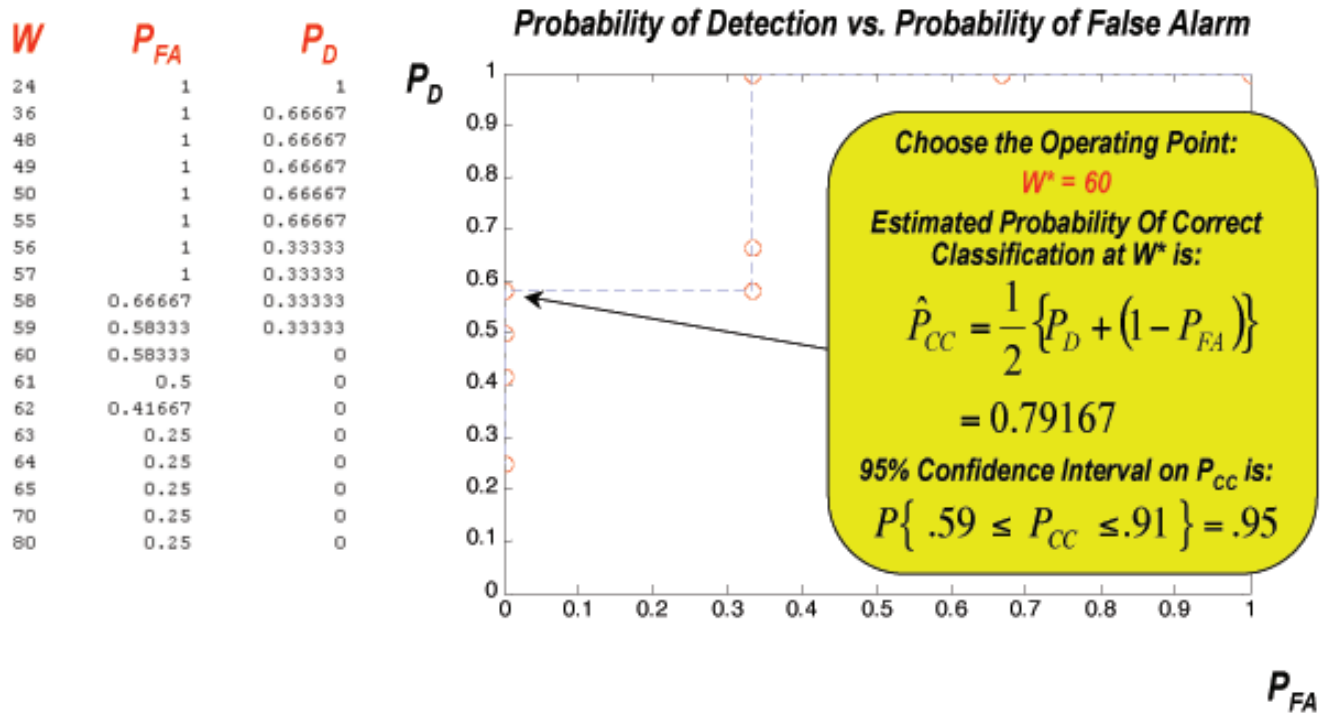


Figure 9.18: E1 All Damage Types Case: The Receiver Operating Characteristic Curve (ROC) yields an operating point with $P(CC) = .792$. We note that because the statistical sample size is small, the confidence interval on the probability of correct classification is wide.

Chapter 10

The Final TDR Data Set

This chapter summarizes the final TDR data set in which we have signals for undamaged cables and for damaged cables. The pulse insertion unit (PIU) and the cable were removed and replaced inside the 2D fixture between data acquisitions. All of the cables in our inventory were measured and the signals were stored for use as exemplars for the undamaged case. The cable damage types were cuts, crushes and pinholes. Cut damage was created using scissors, crush damage was made using a press and steel rods of various sizes, and pinholes were made with a sharp steel punch. The figures in this chapter present representative signals from the set of signals representing cut damage, crush damage and pinhole damage.

10.1 Baseline for the Undamaged Cable Case

A baseline data set for each of our cables was acquired using the TDR 2D fixture and the operating procedure specified in report AF-PT41 issue D. Each cable has a minimum of three waveforms for the undamaged case, while cables that were damaged often have more. The cable was completely removed from the 2D fixture and the PIU between acquisitions for the baseline data set. The files are stored in the folder named "09-23-08_BaselineIII" found on the data CD provided for this project. The file format is "XXXX-baselineYY-ZZ.txt"where "XXXX"denotes the cable serial number, "YY"denotes the baseline study, and "ZZ"denotes the waveform number.

Three baseline studies (III, IV and V) were conducted using the current procedure, AF-PT4194 issue D. Baseline IV and V used a different TDR pulser than Baseline III. Baseline V implements a thirty minute warmup time for the pulser. This warmup time was not used in the other studies. It was intended to improve repeatability.

10.2 Cut Damage Measurements

The cut damage signals are plotted in Figures (10.1), (10.2), (10.3), (10.4), (10.5), (10.6), (10.7), (10.8) and (10.9). The figure captions describe the damage scenarios.

10.3 Pinhole Damage Measurements

Figure (10.10) plots the signal for the pinhole damage case. The cable was shorted at the damage site. Pinhole damage was introduced at around 12 inches from the connector. The damage is to the cable insulation and conductor going all the way through both conductors.

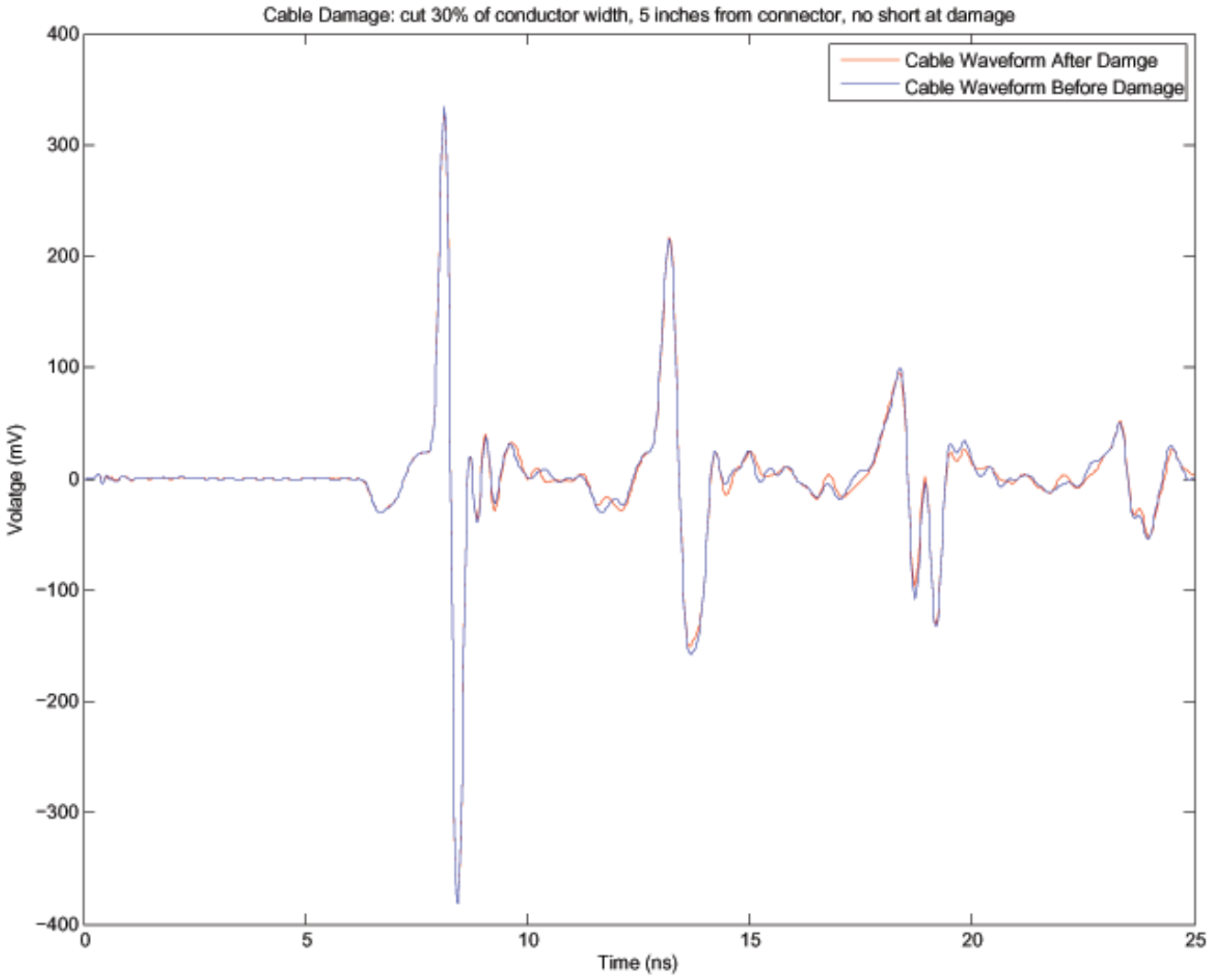


Figure 10.1: **Cut Damge:** A cut was introduced at around 5 inches from the connector. The cut damage is to the cable insulation and conductor cutting through about 30 percent of the conductors width. The conductors are NOT SHORTED at the damage site.

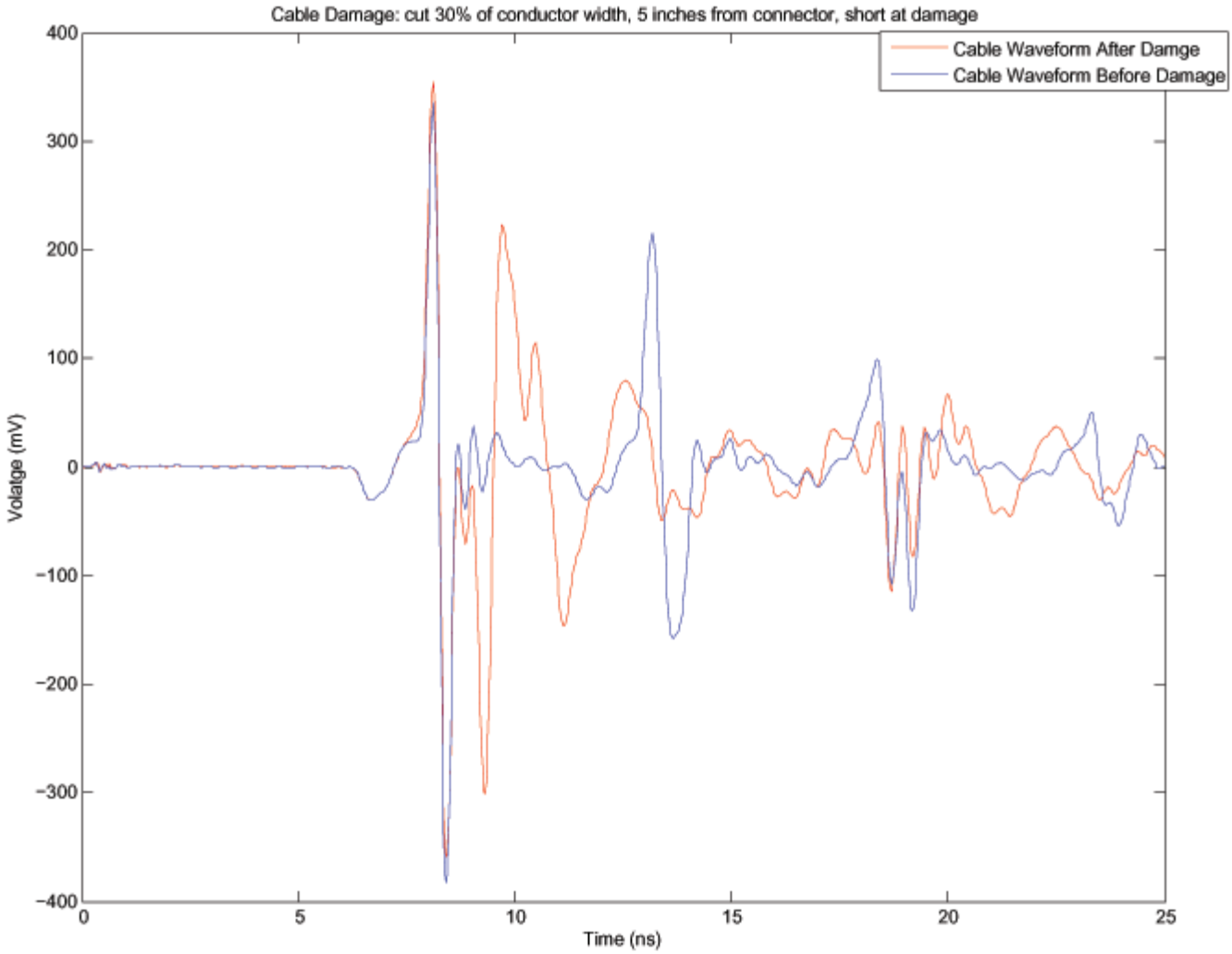


Figure 10.2: **Cut Damage:** A cut was introduced at around 5 inches from the connector. The cut damage is to the cable insulation and conductor cutting through about 30 percent of the conductors width. The conductors ARE SHORTED at the damage site.

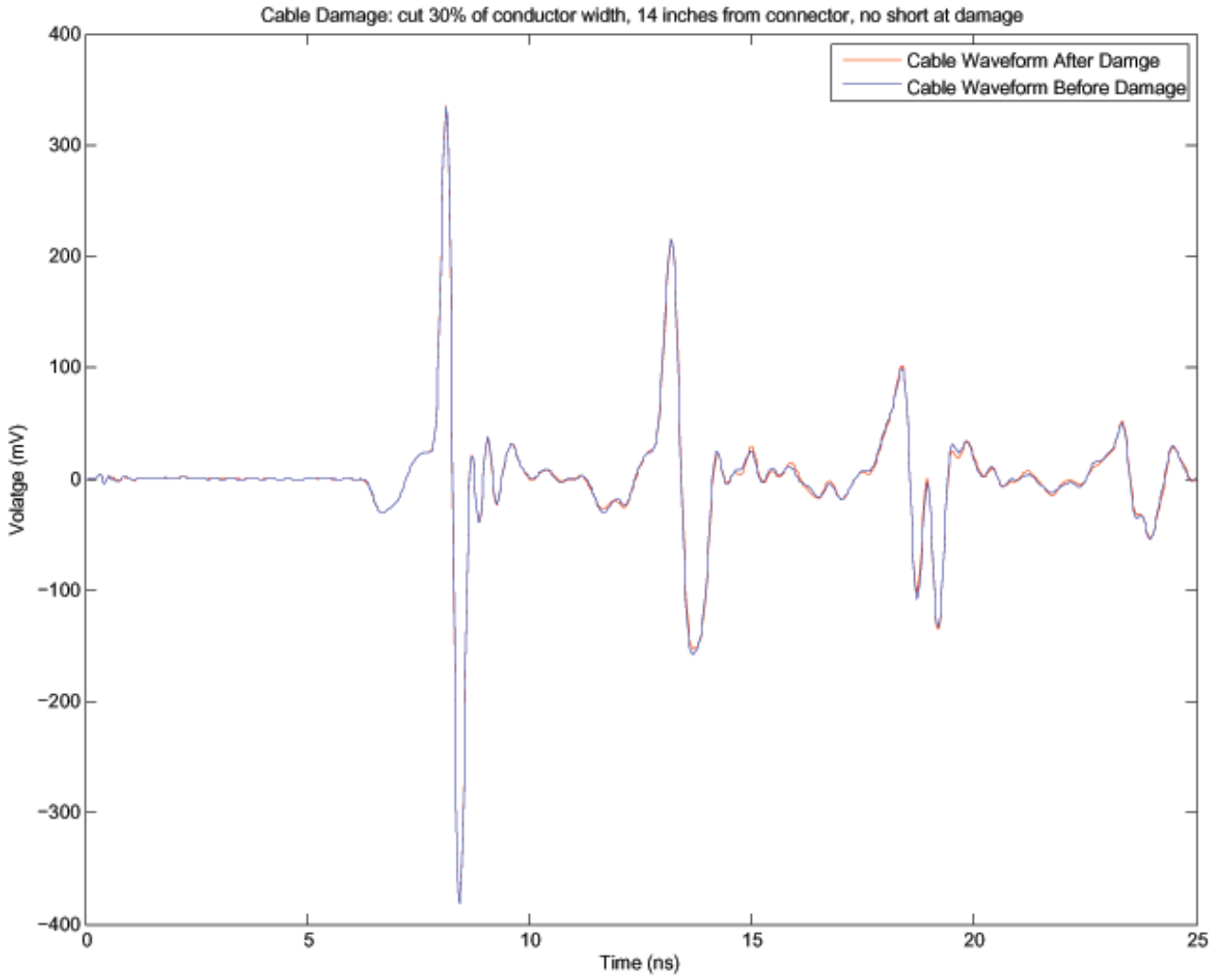


Figure 10.3: **Cut Damage:** A cut was introduced at around 14 inches from the connector. The cut damage is to the cable insulation and conductor cutting through about 30 percent of the conductors width. The conductors are NOT SHORTED at the damage site.

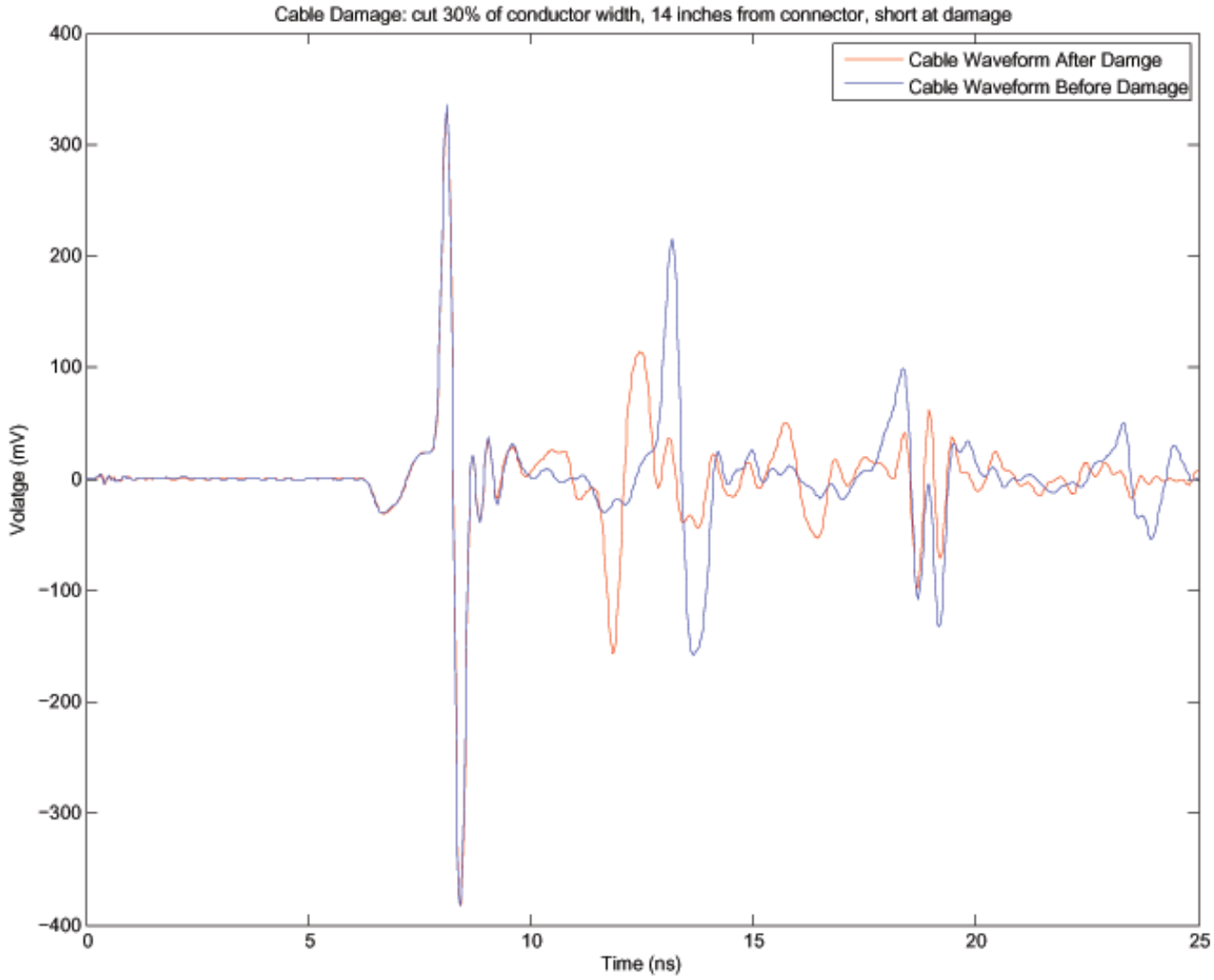


Figure 10.4: **Cut Damage:** A cut was introduced at around 14 inches from the connector. The cut damage is to the cable insulation and conductor cutting through about 30 percent of the conductors width. The conductors ARE SHORTED at the damage site.

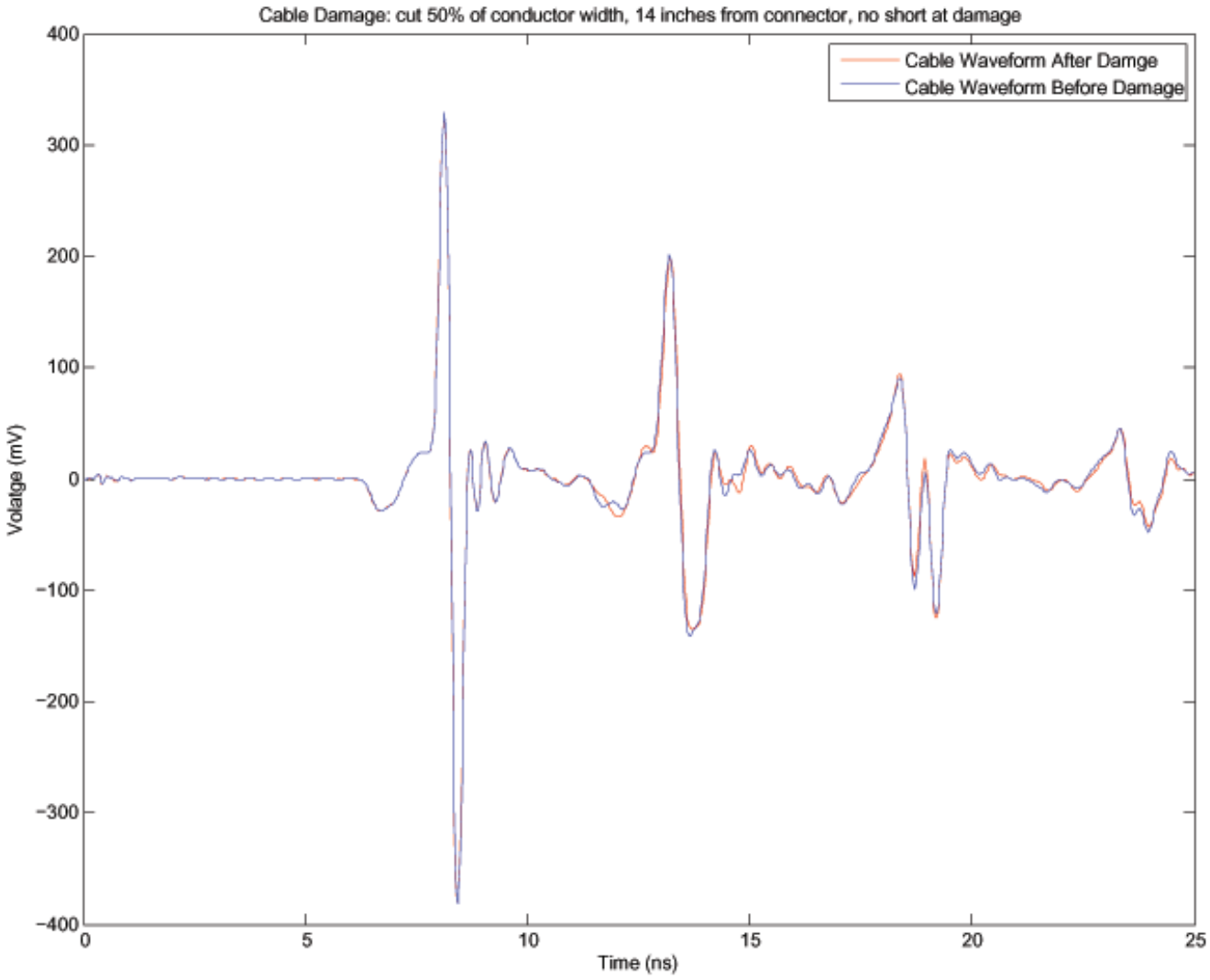


Figure 10.5: **Cut Damage:** A cut was introduced at around 14 inches from the connector. The cut damage is to the cable insulation and conductor cutting through about 50 percent of the conductors width. The conductors are NOT SHORTED at the damage site.

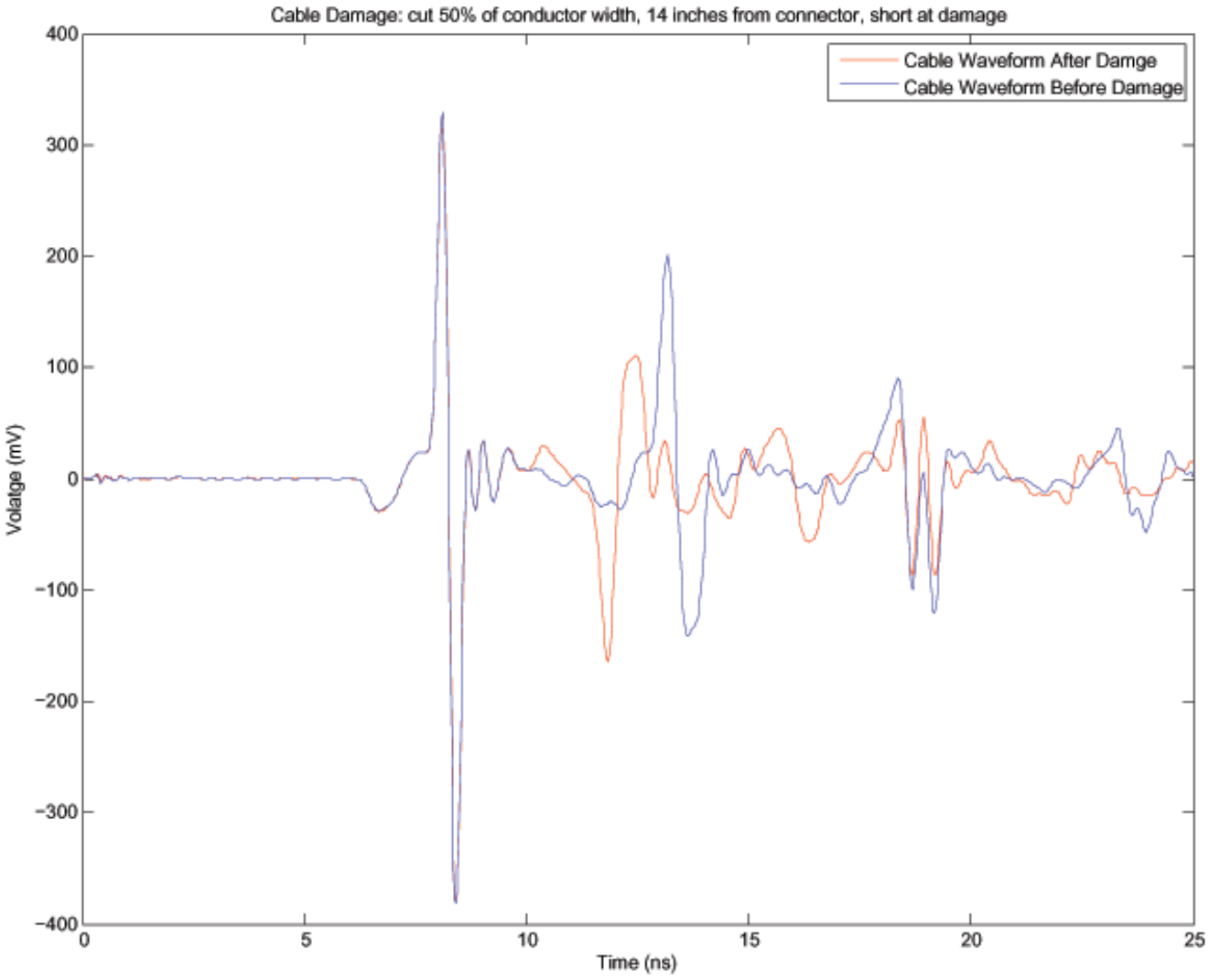


Figure 10.6: **Cut Damage:** A cut was introduced at around 14 inches from the connector. The cut damage is to the cable insulation and conductor cutting through about 50 percent of the conductors width. The conductors ARE SHORTED at the damage site.

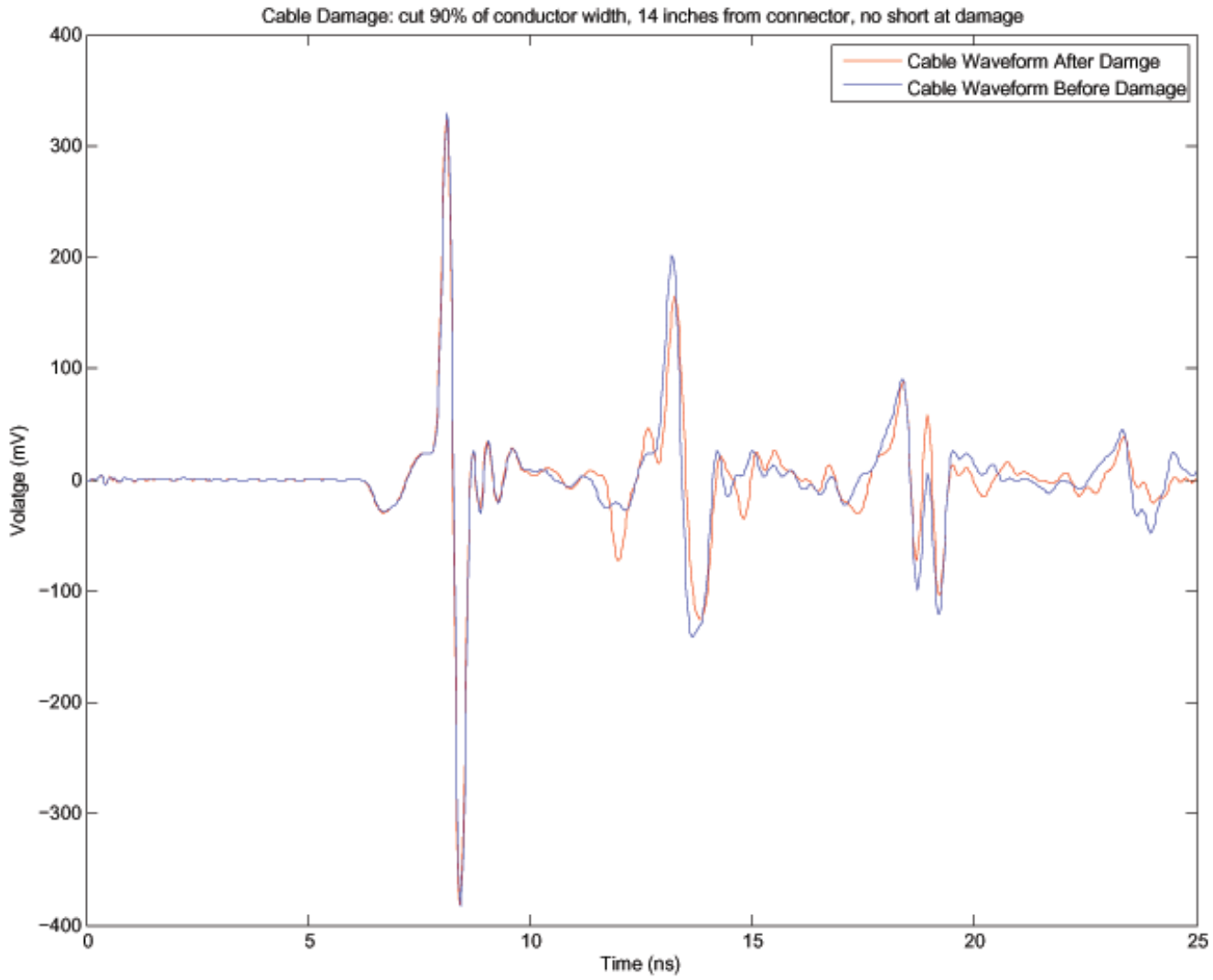


Figure 10.7: **Cut Damage:** A cut was introduced at around 14 inches from the connector. The cut damage is to the cable insulation and conductor cutting through about 90 percent of the conductors width. The conductors are NOT SHORTED at the damage site.

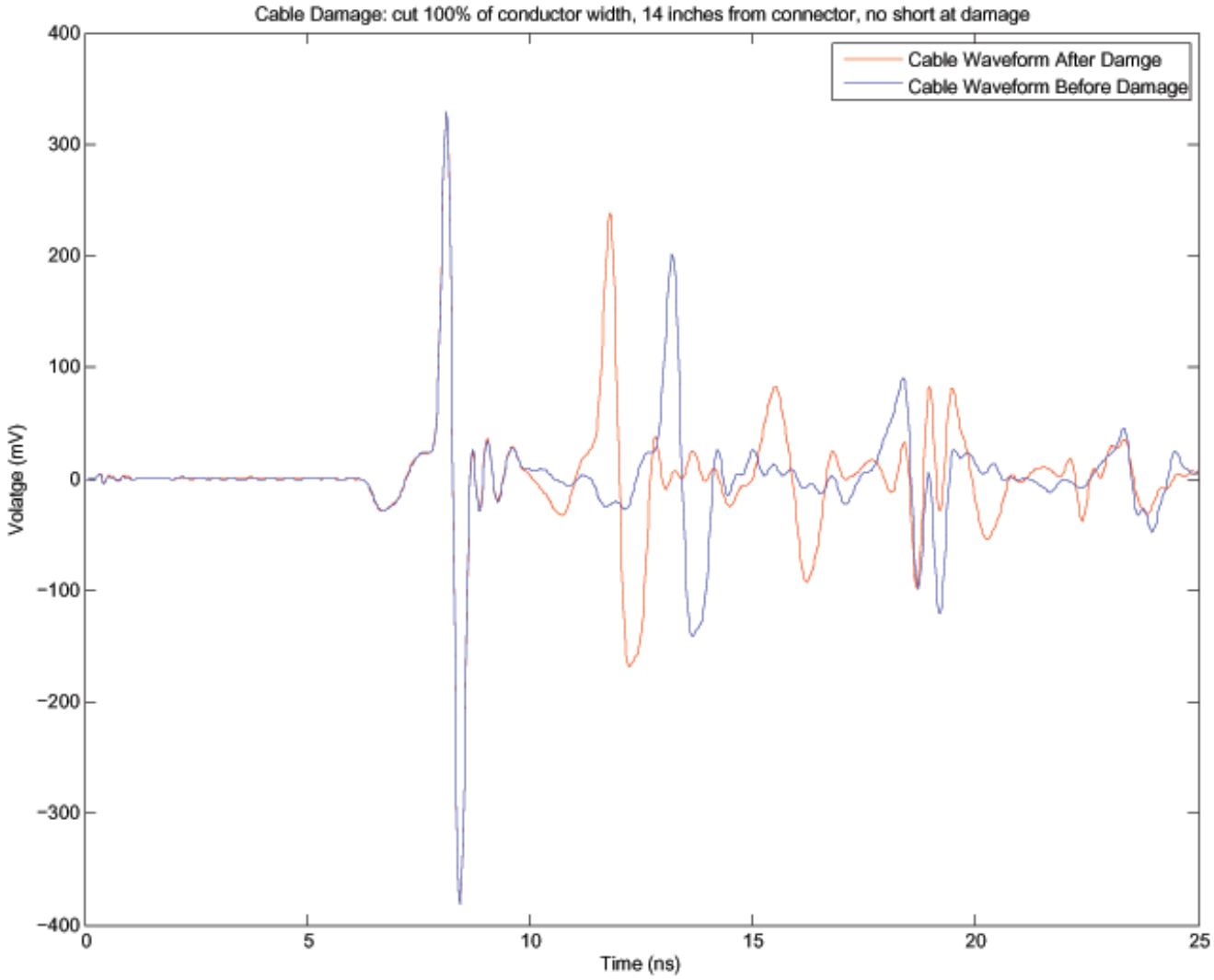


Figure 10.8: **Cut Damage:** A cut was introduced at around 14 inches from the connector. The cut damage is to the cable insulation and conductor cutting all the way through the conductors width. The conductors are NOT SHORTED at the damage site.

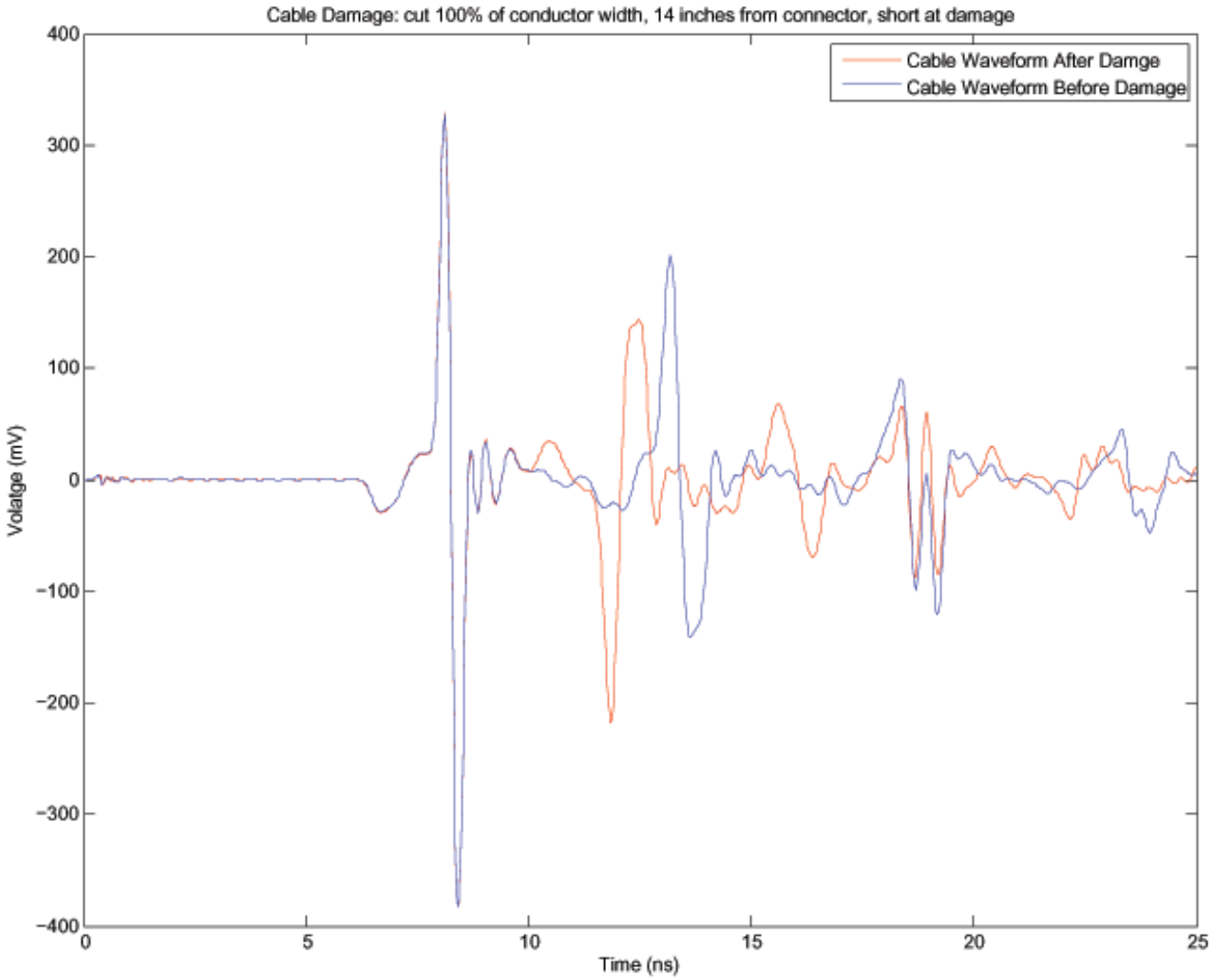


Figure 10.9: **Cut Damage:** A cut was introduced at around 14 inches from the connector. The cut damage is to the cable insulation and conductor cutting all the way through the conductors width. The conductors ARE SHORTED at the damage site.

We have sets of pinhole damage results for two different cables. These resulted in both shorted and not shorted conductors. These results are not included in this report, but they are available in the data archives.

10.4 Crush Damage Measurements

The crush damage signals are plotted in Figures (10.11) and (10.12). The figure captions describe the damage scenarios. It should be noted that while the TDR cannot detect crush damage, all cables damaged in this manner successfully passed high pot testing. Therefore, these cables cannot be truly damaged from crush insults until the crush is sufficient to split the cable in two.

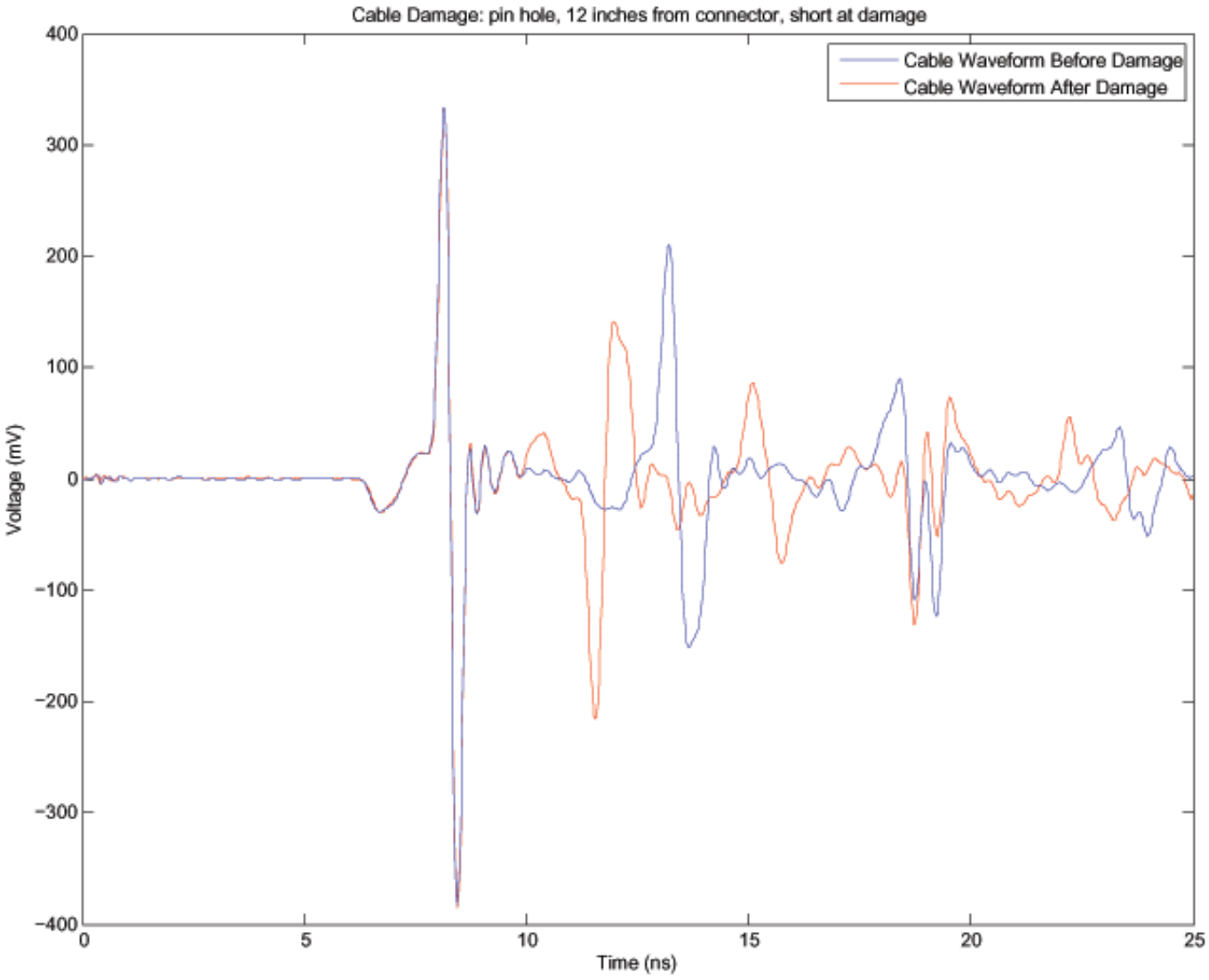


Figure 10.10: **PinholeDamage at 12 in., short:** Pin-hole damage was introduced at around 12 inches from the connector. The damage is to the cable insulation and conductor going all the way through both conductors. The conductors ARE SHORTED at the damage site.

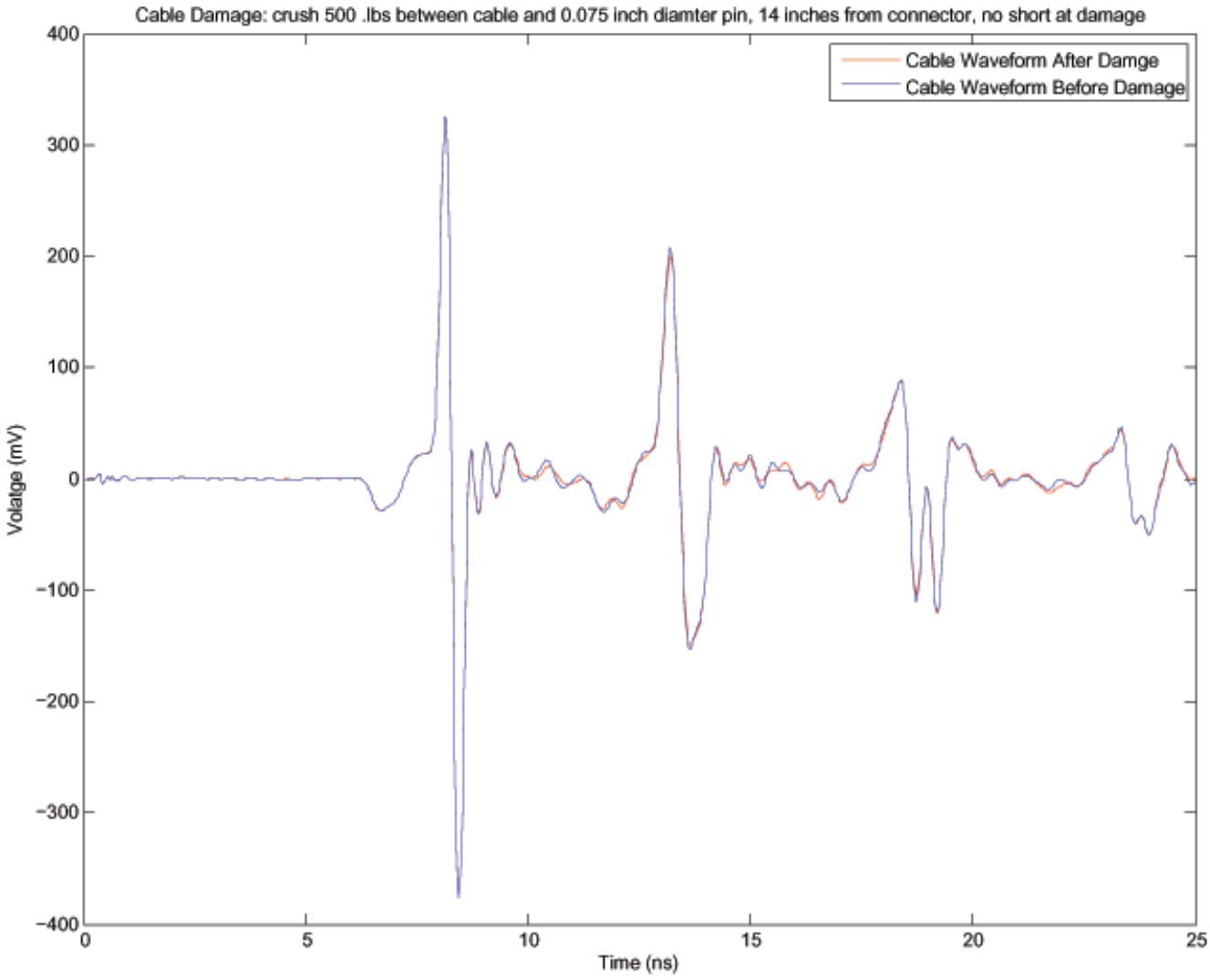


Figure 10.11: **Crush Damage:** A crush was introduced at around 14 inches from the connector. The crush damage is to the cable insulation and conductor by pressing a 0.075 inch diameter pin across the cable width at 500 lbs of force. The conductors are NOT SHORTED at the damage site and withstood a high-pot test to 7kV.

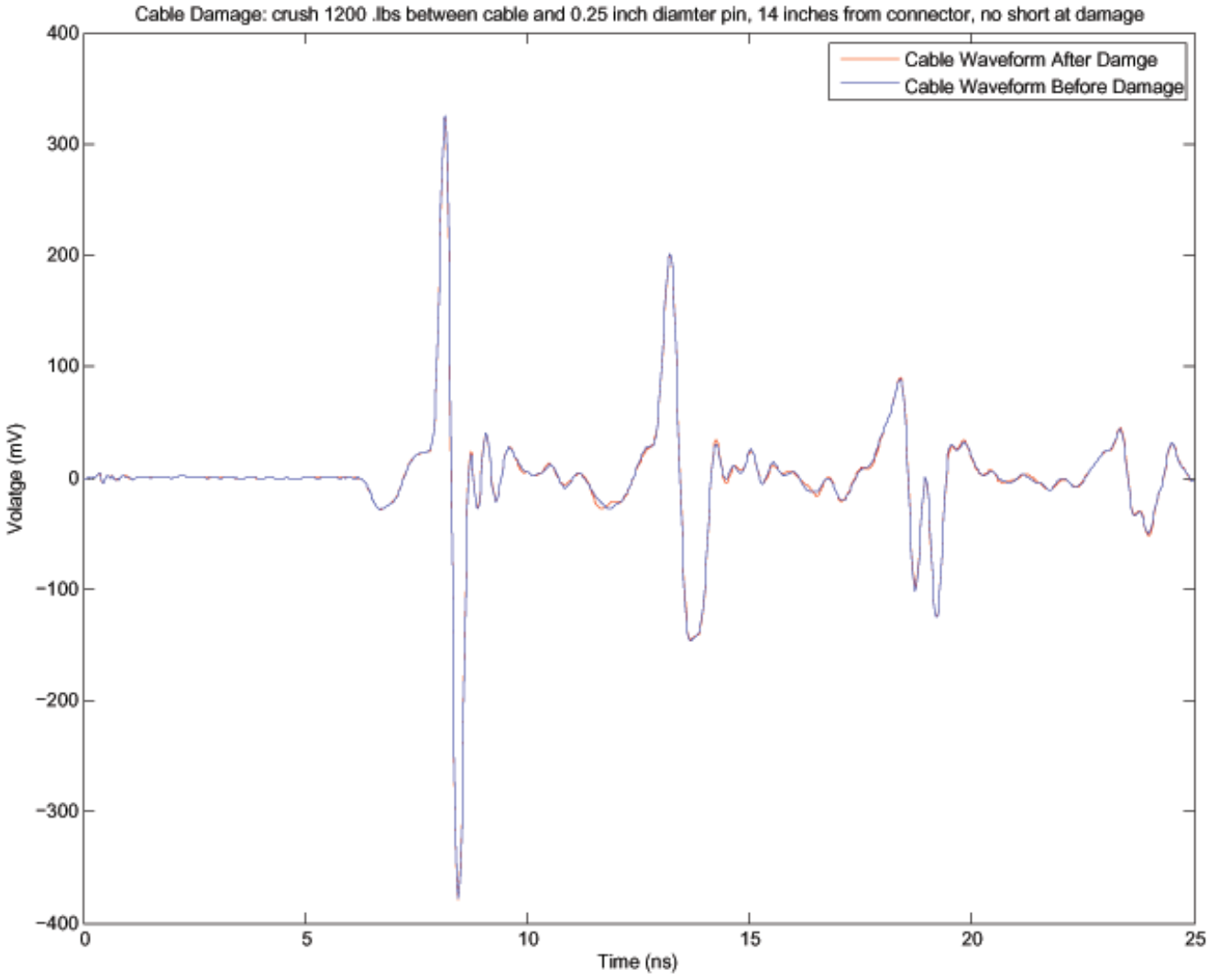


Figure 10.12: **Crush Damage:** A crush was introduced at around 14 inches from the connector. The crush damage is to the cable insulation and conductor by pressing a 0.250 inch diameter pin across the cable width at 1200 lbs of force. The conductors are NOT SHORTED at the damage site and withstood a high-pot test to 7kV.

Chapter 11

Experiment E5: Analysis of the Final TDR Data Set - "Insult " Types are Cuts and Crushes

This section represents final results, in the sense that it used the final data set to assess the experimental procedures, hardware, software, etc. As described in the previous chapter, the pulse insertion unit (PIU) and the cable were removed and replaced inside the 2D fixture between data acquisitions. All of the cables in our inventory were measured and the signals were stored for use as exemplars for the undamaged case. The cable damage types were cuts, crushes and pinholes. Cut damage was created using scissors, crush damage was made using a press and steel rods of various sizes, and pinholes were made with a sharp steel punch. The figures in this chapter present representative signals from the set of signals representing cut damage, crush damage and pinhole damage.

A baseline data set for each of our cables was acquired using the TDR 2D fixture and the operating procedure specified in report AF-PT41 issue D. Each cable has a minimum of three waveforms for the undamaged case, while cables that were damaged often have more. The cable was completely removed from the 2D fixture and the PIU between acquisitions for the baseline data set. The files are stored in the folder named "09-23-08_BaselineIII" found on the data CD provided for this project. The file format is "XXXX-baselineYY-ZZ.txt"where "XXXX"denotes the cable serial number, "YY"denotes the baseline study, and "ZZ"denotes the waveform number.

Three baseline studies (III, IV and V) were conducted using the current procedure, AF-PT4194 issue D. Baseline IV and V used a different TDR pulser than Baseline III. Baseline V implements a thirty minute warmup time for the pulser. This warmup time was not used in the other studies. It was intended to improve repeatability.

The ARX model identified for the baseline case and used for model-based damage detection is the following: $ARX(N_a, N_b, N_k) = ARX(20, 20, 3)$.

Note that the ROC Curves shown in this chapter are not smooth because we had only six sample points with which to estimate the P_D and P_{FA}. The number six comes from the fact that we had available only three undamaged cable signals and three damaged cable signals. Only the cut cables were analyzed with the signal processing algorithms. The pinhole damage was very similar to the cut data, and the crush damage turned out to be superficial, as all cables with crush damage passed the required hi pot test.

Cables were cut with scissors from the side. The cables and the cuts are summarized as follows:

Case (1) Damage Resulting in a Short Circuit:
1403-cut-30percent-short, 3,4,5: TDR Birth Certificate, No TDR Birth Certificate Cases
The way to read this title is the following: Cable number 1403 was cut from the side through 30 percent of the

width of the cable at approximately 14 inches from the connector. The cut resulted in a short circuit. The signals used for analysis were the third, fourth and fifth signals measured for this cable scenario. Two cases were studied: (a) The case in which the cable had a "TDR birth certificate "and (b) The case in which the cable did not have a "TDR birth certificate. " Having a birth certificate means that the signals for the undamaged case came from cable 1403 before it was damaged. Having no TDR birth certificate means that the signals for the undamaged case came from a pool of signals measured for other undamaged cables.

This case corresponds to the damage represented in Figure (10.4). One can see from the figure that the damage is extremely obvious by visual inspection of the signal. This gives us a chance to test the algorithms on a case that should clearly give detections. We see from the results presented below that the damage detection algorithms perform ideally for this case.

Case (2) Damage Not Resulting in a Short Circuit:
1422-cut-90percent-noshort,7,8,9: TDR Birth Certificate, No TDR Birth Certificate Cases

The way to read this title is the following: Cable number 1422 was cut from the side through 90 percent of the width of the cable at approximately 14 inches from the connector. The cut did not result in a short circuit. The signals used for analysis were the seventh, eigth and ninth signals measured for this cable scenario. Two cases were studied: (a) The case in which the cable had a "TDR birth certificate "and (b) The case in which the cable did not have a "TDR birth certificate. " Having a birth certificate means that the signals for the undamaged case came from cable 1422 before it was damaged. Having no TDR birth certificate means that the signals for the undamaged case came from a pool of signals measured for other undamaged cables.

This case corresponds to the damage represented in Figure (10.7). One can see from the figure that the damage is reasonably obvious by visual inspection of the signal. The damage signal corresponds to the undamaged signal reasonably well until about the time that the back reflection arrives - then it departs significantly. However, the effects of the damage are not as severe as they are in case (1) above. We see from the results presented below that for this case, the damage detection performance is less than ideal. We expected better performance. Future plans include reworking this experiment, explaining the suboptimal performance and making improvements as necessary.

Case (3) Damage Not Resulting in a Short Circuit:
1422-cut-50percent-noshort,1,2,3: TDR Birth Certificate, No TDR Birth Certificate Cases

The way to read this title is the following: Cable number 1422 was cut from the side through 50 percent of the width of the cable. The cut did not result in a short circuit. The signals used for analysis were the first, second and third signals measured for this cable scenario. Two cases were studied: (a) The case in which the cable had a "TDR birth certificate "and (b) The case in which the cable did not have a "TDR birth certificate. " Having a TDR birth certificate means that the signals for the undamaged case came from cable 1422 before it was damaged. Having no TDR birth certificate means that the signals for the undamaged case came from a pool of signals measured for other undamaged cables.

This case corresponds to the damage represented in Figure (10.5). One can see from the figure that the damage appears subtle by visual inspection of the signal. The damage signal corresponds to the undamaged signal reasonably well until about the time that the back reflection arrives - then it departs somewhat. The effects of the damage are somewhat subtle, and not as nearly as severe as they are in case (1) or case (2) above. We see from the results presented below that for this case, the damage detection performance is very poor. Future plans include reworking this experiment, explaining the suboptimal performance and making improvements as necessary.

11.1 Case (1) with No TDR Birth Certificate Results: 1403-cut-30percent-short, 3,4,5, No TDR Birth Certificate

For this case, the ROC curve is shown in Figure (11.1). The confidence interval bounds on the probability of correct classification are plotted in Figure (11.2). The confidence intervals for the sample size used ($N = 6$) and

also for the cases in which $N = 100$ and $N = 1000$ for reference. They are shown for both the Gaussian and exact cases in Figure (11.3).

This case corresponds to the damage represented in Figure (10.4). One can see from the figure that the damage is extremely obvious by visual inspection of the signal. This gives us a chance to test the algorithms on a case that should clearly give detections.

We see from the results presented below that the damage detection algorithms perform ideally for this case.

11.2 Case (1) with TDR Birth Certificate Results: 1403-cut-30percent-short, 3,4,5, TDR Birth Certificate Case

For this case, the ROC curve is shown in Figure (11.4). The confidence interval bounds are plotted in Figure (11.5).

11.3 Case (2) with No TDR Birth Certficate Results: 1422-cut-90percent-noshort,7,8,9, No Birth Certificate Case

For this case, the ROC curve is shown in Figure (11.6). The confidence interval bounds are plotted in Figure (11.7).

This case corresponds to the damage represented in Figure (10.7). One can see from the figure that the damage is reasonably obvious by visual inspection of the signal. The damage signal corresponds to the undamaged signal reasonably well until about the time that the second reflection arrives - then it departs significantly. However, the effects of the damage are not as severe as they are in case (1) above. We see from the results presented below that for this case, the damage detection performance is less than ideal. Future plans include reworking this experiment, explaining the suboptimal performance and making improvements as necessary.

11.4 Case (2) with TDR Birth Certificate Results: 1422-cut-90percent-noshort,7,8,9, Birth Certificate Case

For this case, the ROC curve is shown in Figure (11.8). The confidence interval bounds are plotted in Figure (11.9).

11.5 Case (3) with No TDR Birth Certificate Results: 1422-cut-50percent-noshort,1,2,3, No Birth Certificate Case

For this case, the ROC curve is shown in Figure (11.10). The confidence interval bounds are plotted in Figure (11.11).

This case corresponds to the damage represented in Figure (10.5). One can see from the figure that the damage is not very obvious by visual inspection of the signal. The damage signal corresponds to the undamaged signal reasonably well until about the time that the second reflection arrives - then it departs somewhat. The effects of the damage are somewhat subtle, and not as nearly as severe as they are in case (1) or case (2) above. We see from the results presented below that for this case, the damage detection performance is very poor. Future plans include reworking this experiment, explaining the suboptimal performance and making improvements as necessary.

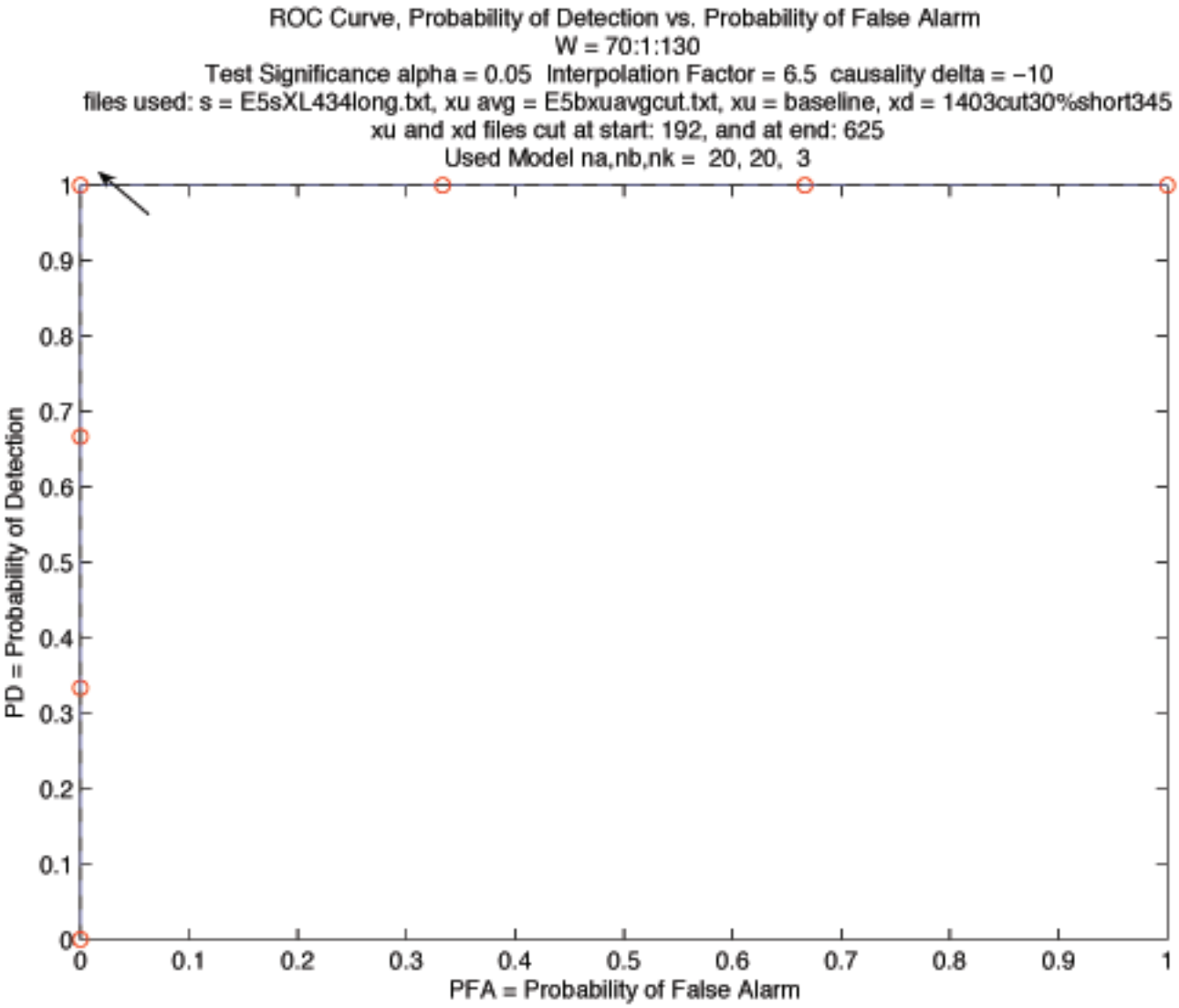


Figure 11.1: **Case (1) ROC Curve, E5b 1403-cut-30percent-short, 3,4,5, No Birth Certificate Case:** The detection performance is ideal for this case. This performance is to be expected, because one can see from Figure (10.4) that the damage is extremely obvious by visual inspection of the signal.

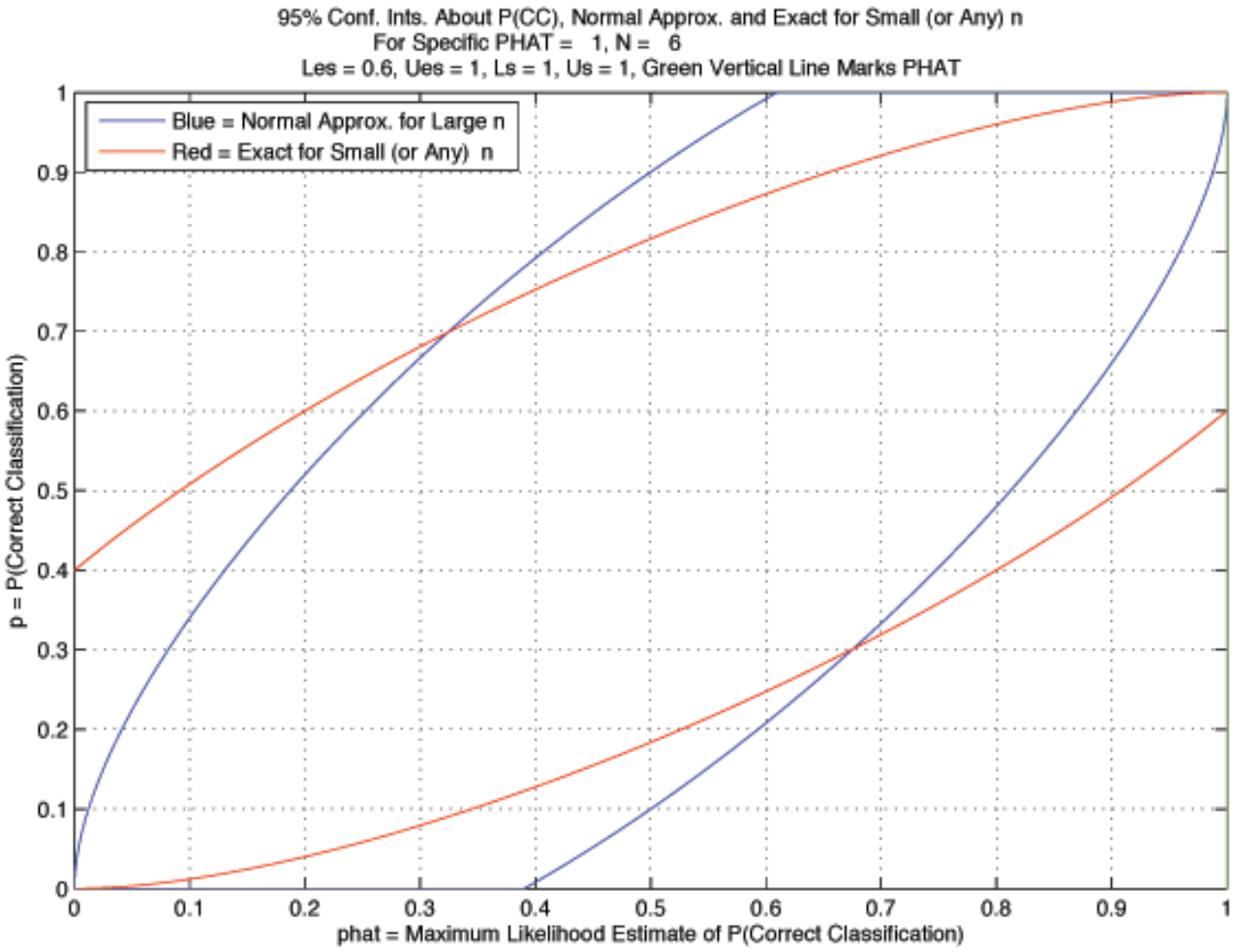


Figure 11.2: **Case (1) with No TDR Birth Certificate: E5b CI Both, 1403-cut-30percent-short, 3,4,5, No Birth Certificate Case:** The confidence interval plots are shown for both the Gaussian and exact cases. The green vertical line marks the calculated estimate of the probability of correct classication $\widehat{p} = 1$.

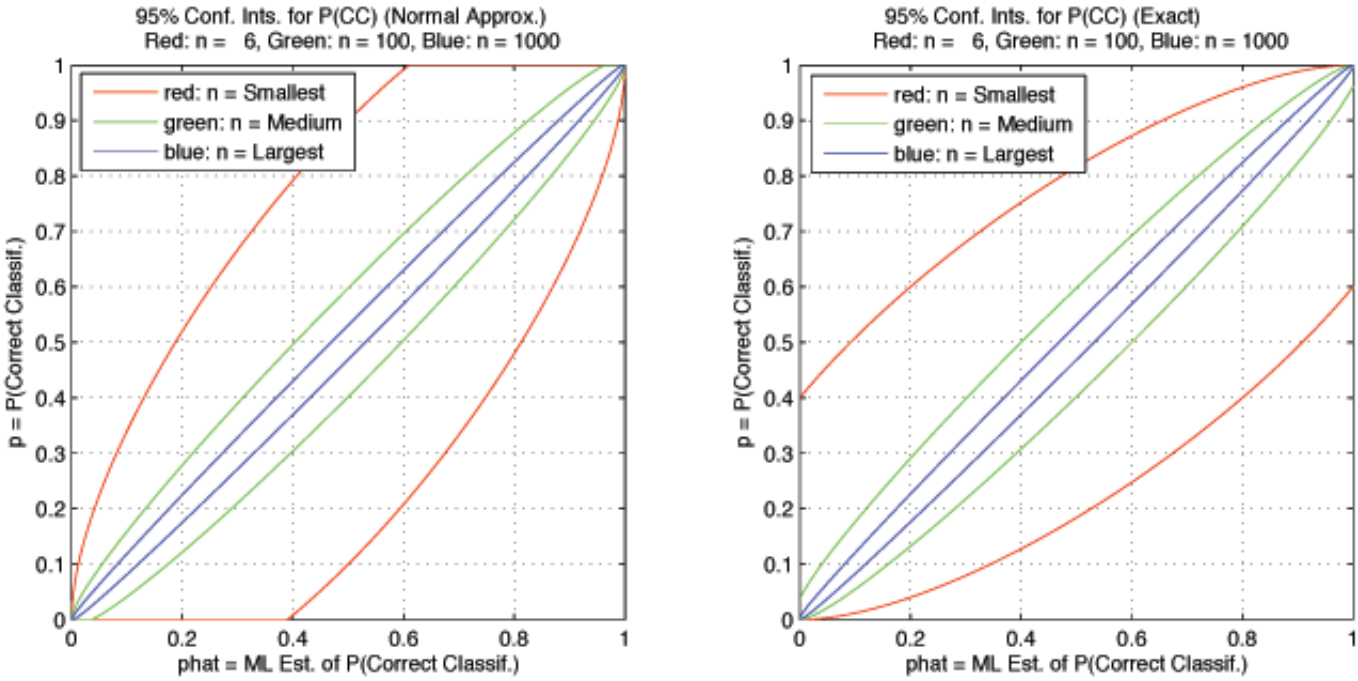


Figure 11.3: **Case (1) with No TDR Birth Certificate: E5b CI All, 1403-cut-30percent-short, 3,4,5, No Birth Certificate Case:** The confidence interval plots for the sample size used ($N = 6$) and also for the cases in which $N = 100$ and $N = 1000$ for reference. They are shown for both the Gaussian and exact cases.

11.6 Case (3) with TDR Birth Certificate Results: 1422-cut-50percent-noshort,1,2,3, TDR Birth Certificate Case

For this case, the ROC curve is shown in Figure (11.12). The confidence interval bounds are plotted in Figure (11.13).

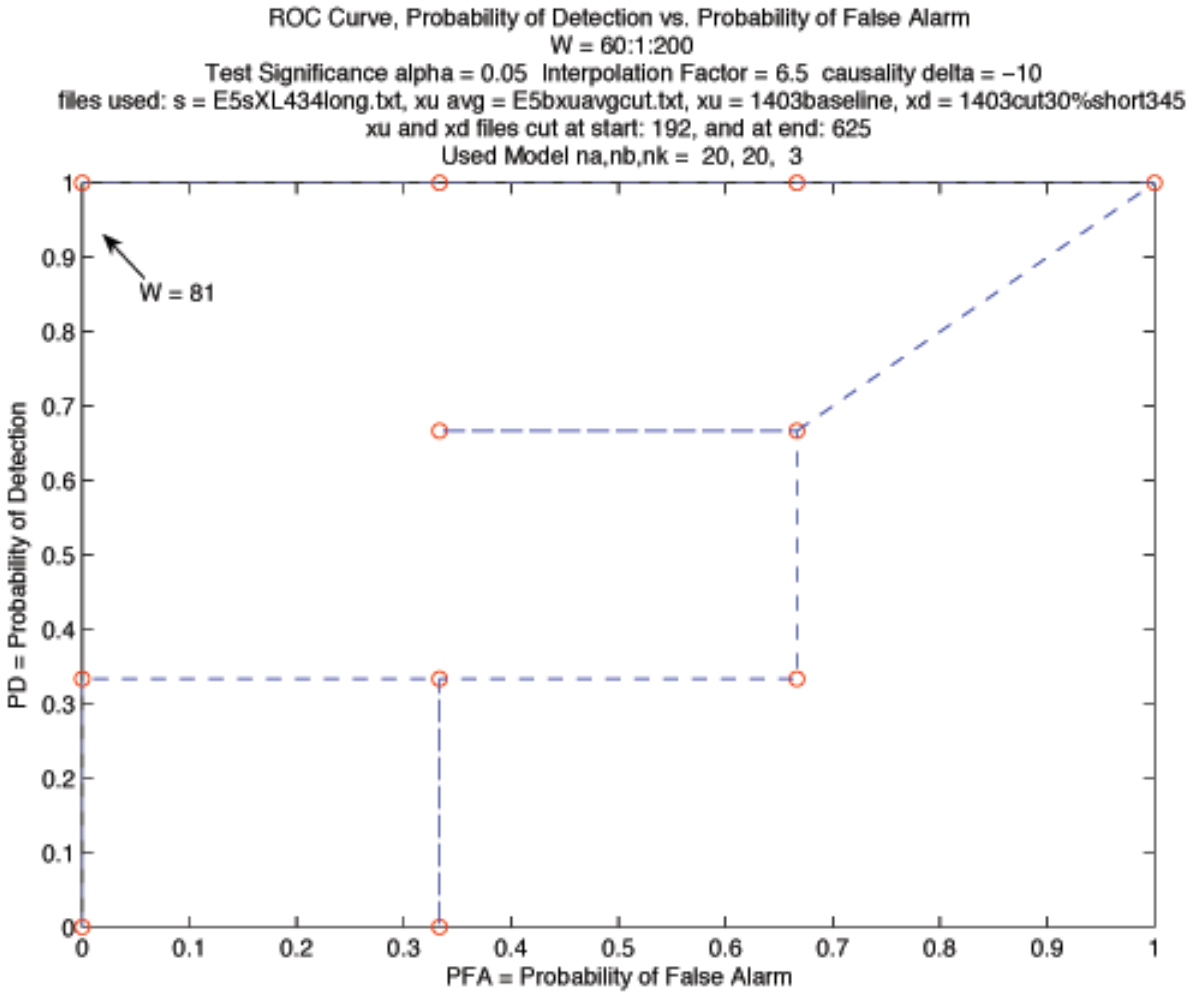


Figure 11.4: **ROC Curve for Case (1) with TDR Birth Certificate Results: E5b ROC Curve, 1403-cut-30percent-short, 3,4,5, Birth Certificate Case:** The ROC Curve is extremely bizarre-looking, and the authors are unable to explain exactly why it is so. The detection performance is ideal for this case, if one picks the ROC operating point in the corner. Future plans include reworking this experiment, explaining the suboptimal performance and making improvements as necessary.

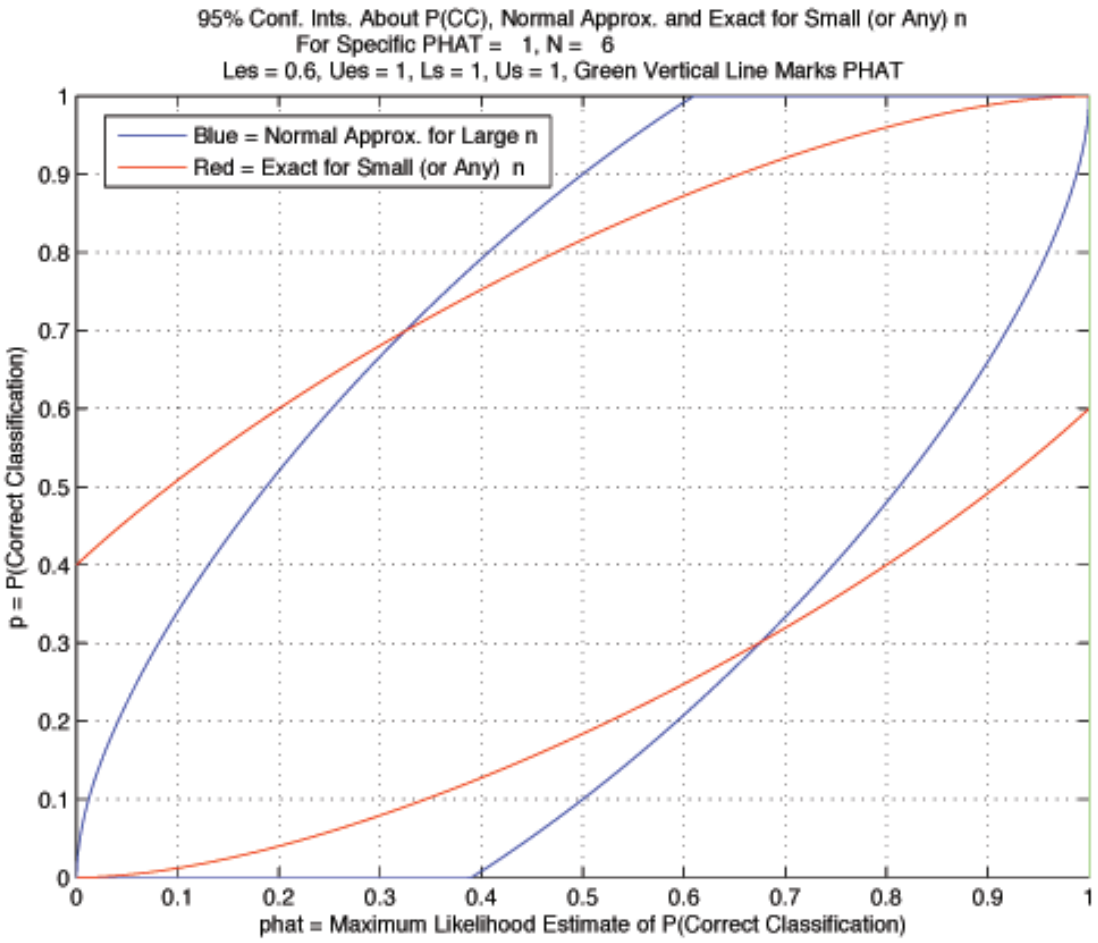


Figure 11.5: **Confidence Interval for Case (1) with TDR Birth Certificate Results: E5b CI Both, 1403-cut-30percent-short, 3,4,5, TDR Birth Certificate Case:** The confidence interval plots are shown for both the Gaussian and exact cases. The green vertical line marks the calculated estimate of the probability of correct classication for the ROC operating point corresponding to $W = 81$ and $\widehat{p} = 1$.

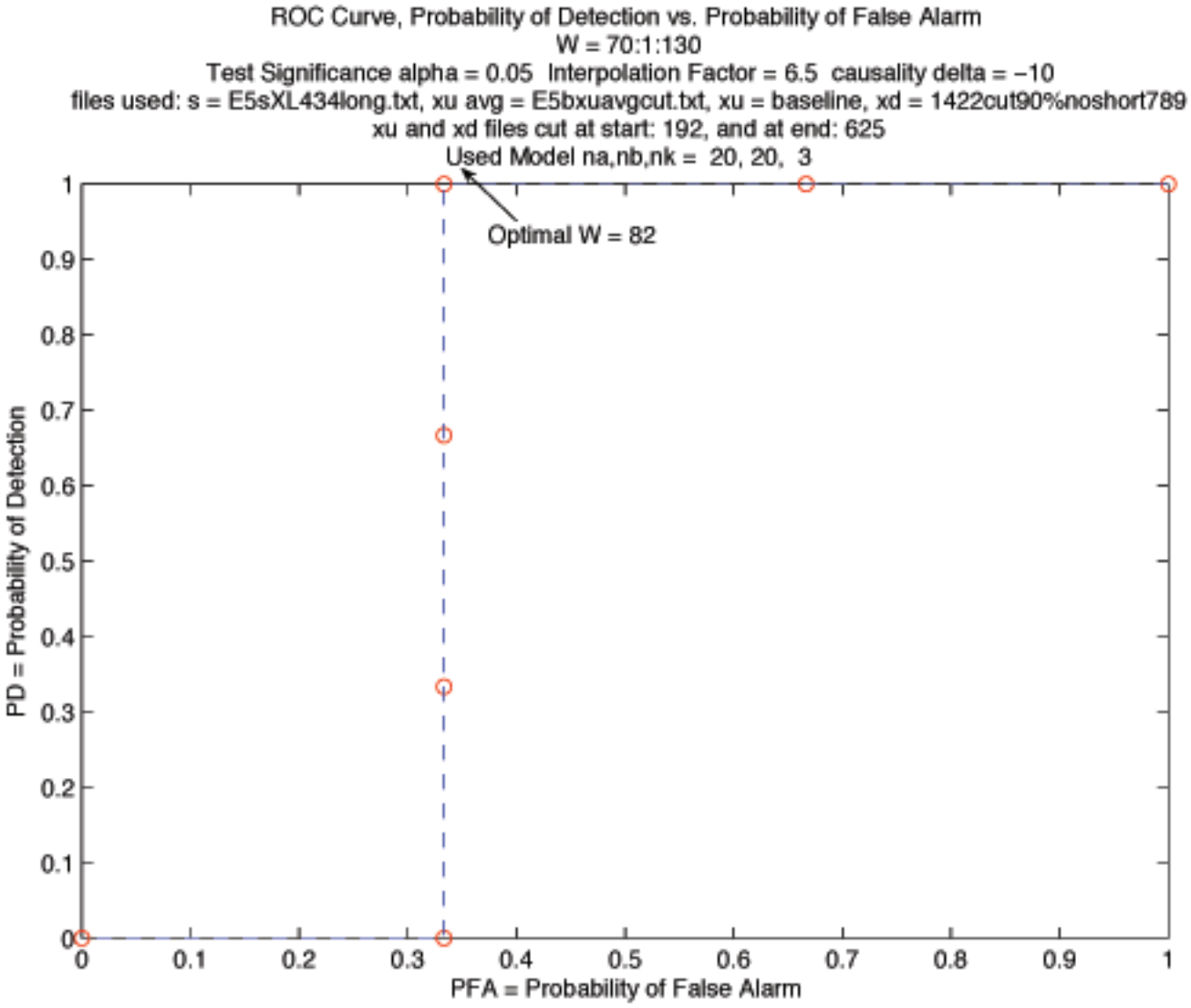


Figure 11.6: **ROC Curve for Case (2) without TDR Birth Certificate Results: E5b, 1422-cut-90percent-noshort,7,8,9, No Birth Certificate Case:** The detection performance is not ideal. Future plans include reworking this experiment, explaining the suboptimal performance and making improvements as necessary.

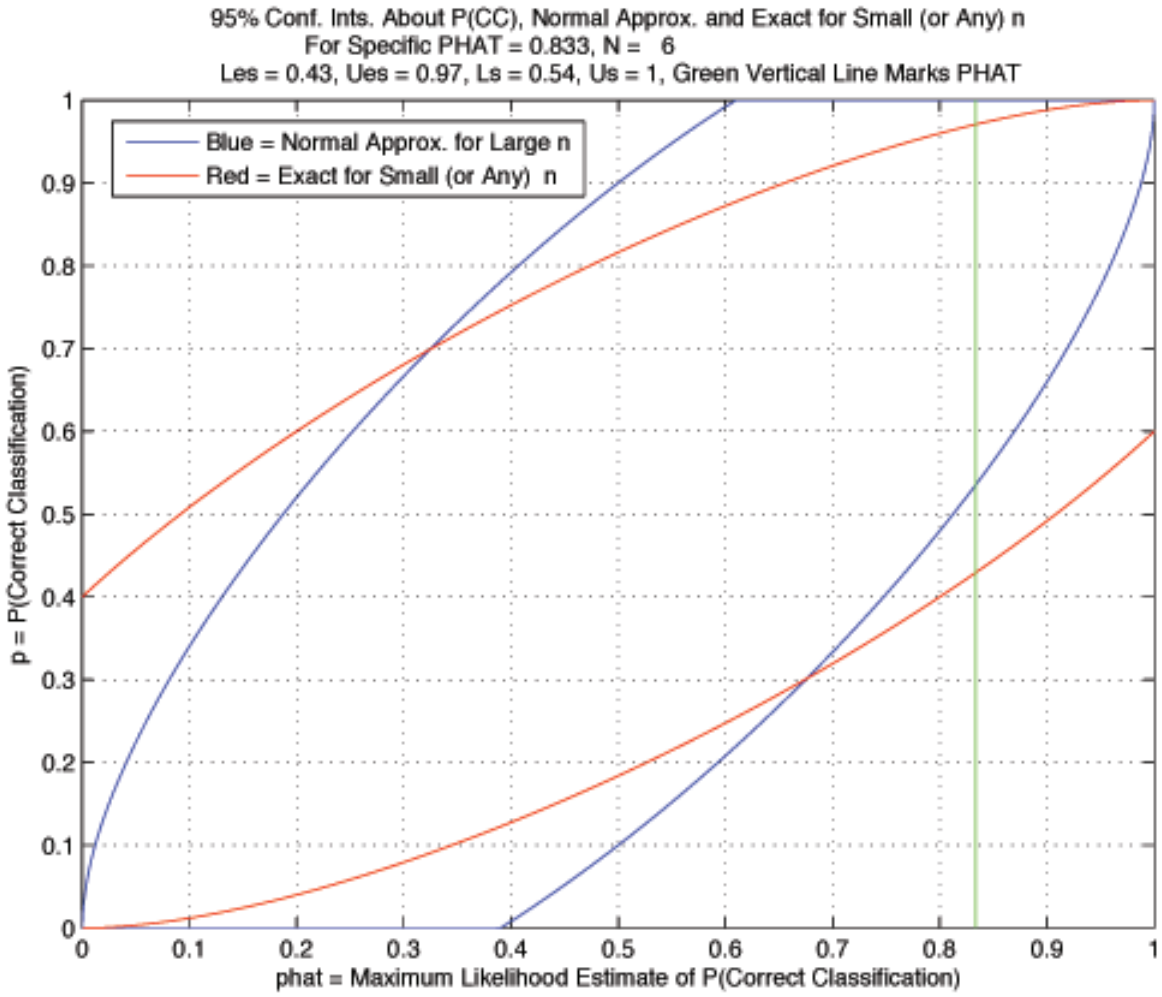


Figure 11.7: **Confidence Intervals for Case (2) with No TDR Birth Certificate Results: E5b CI Both, 1403-cut-30percent-short, 3,4,5, No TDR Birth Certificate Case:** The confidence interval plots are shown for both the Gaussian and exact cases. The green vertical line marks the calculated estimate of the probability of correct classication for the ROC operating point corresponding to $W = 82$ and $\widehat{p} = .833$.

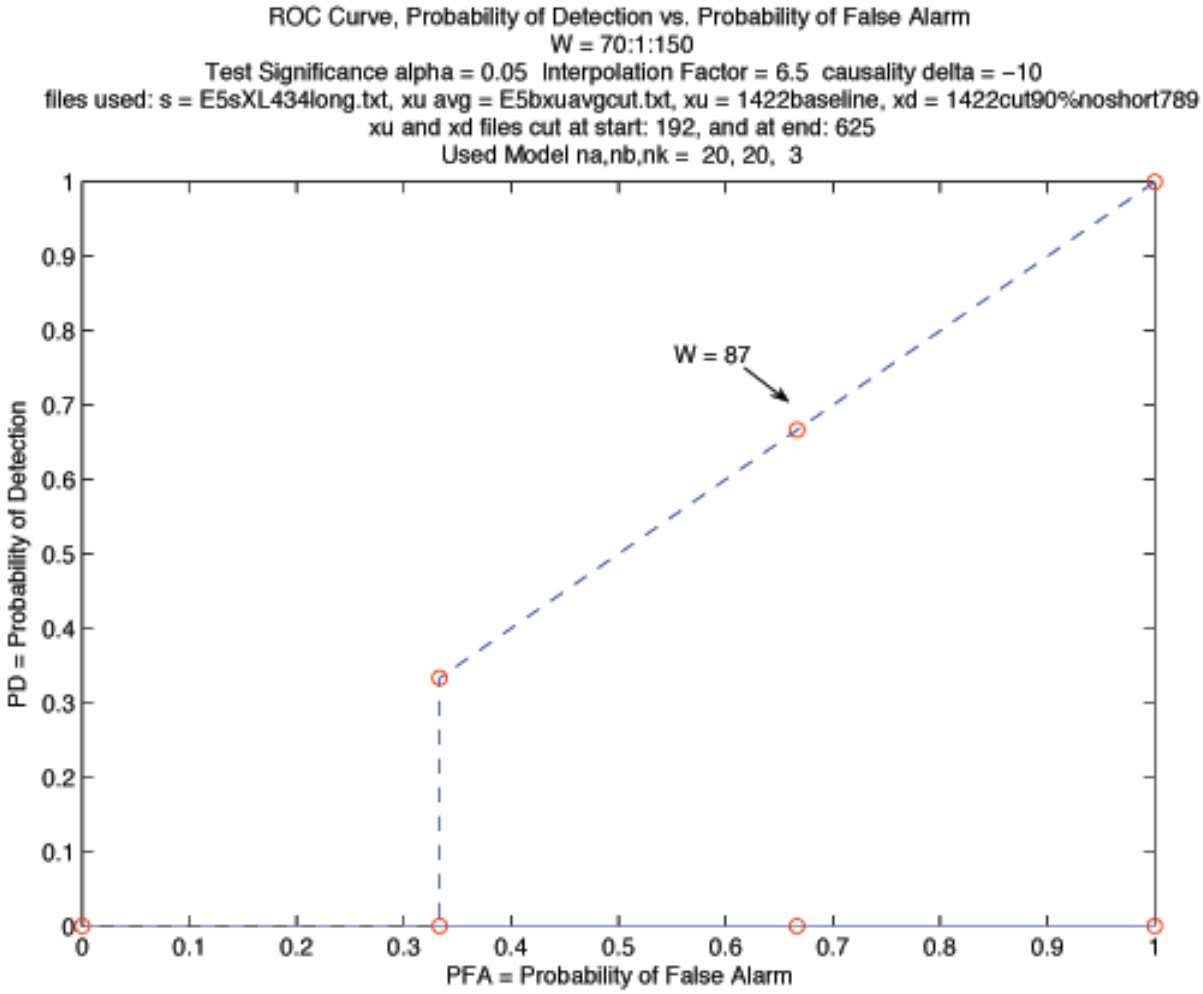


Figure 11.8: **ROC Curve for Case (2) with TDR Birth Certificate Results: E5b ROC Curve, 1422-cut-90percent-noshort,7,8,9, TDR Birth Certificate Case:** The detection performance is very poor. One could flip a coin and do as well. Future plans include reworking this experiment, explaining the poor performance and making improvements as necessary.

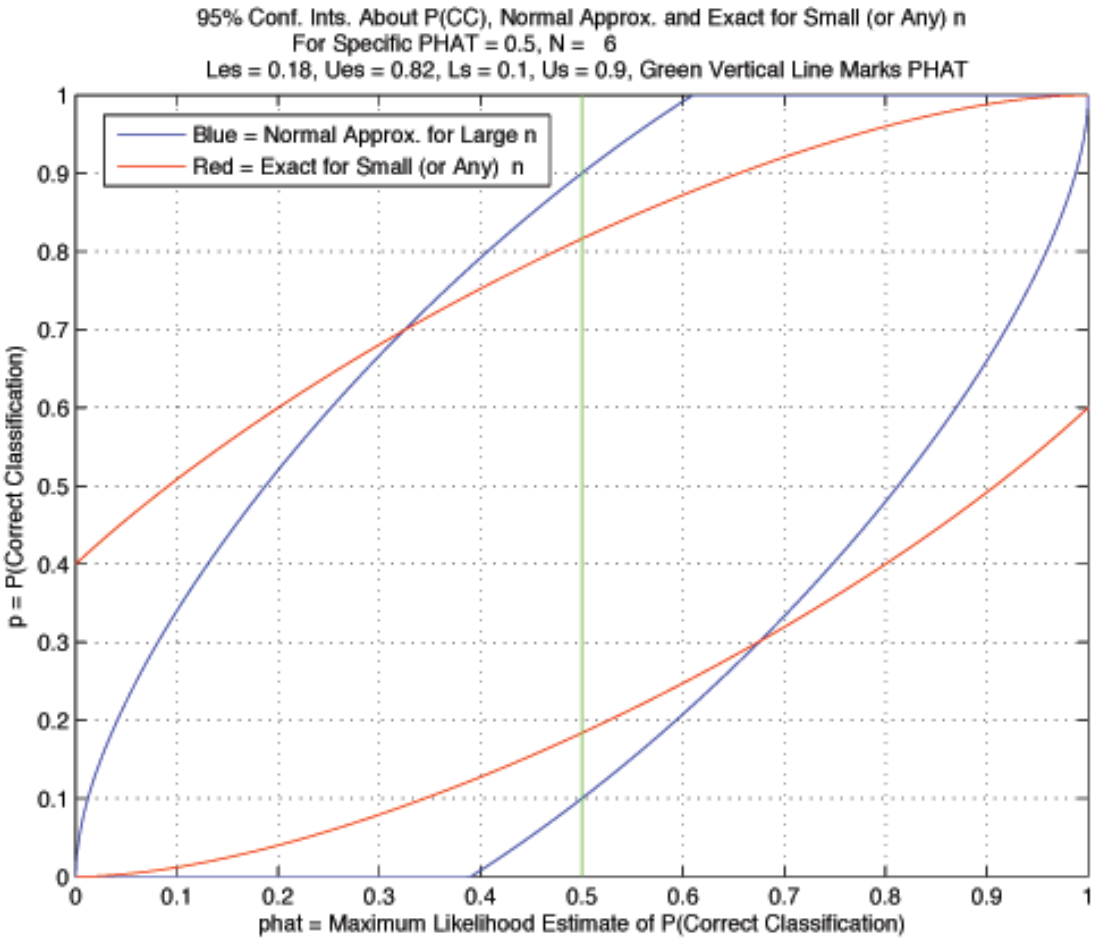


Figure 11.9: **Confidence Interval for Case (2) with TDR Birth Certificate Results: E5b CI Both, 1422-cut-90percent-noshort,7,8,9, TDR Birth Certificate Case:** The confidence interval plots are shown for both the Gaussian and exact cases. The green vertical line marks the calculated estimate of the probability of correct classication for the ROC operating point corresponding to $W = 87$ and $\widehat{p} = .5$.

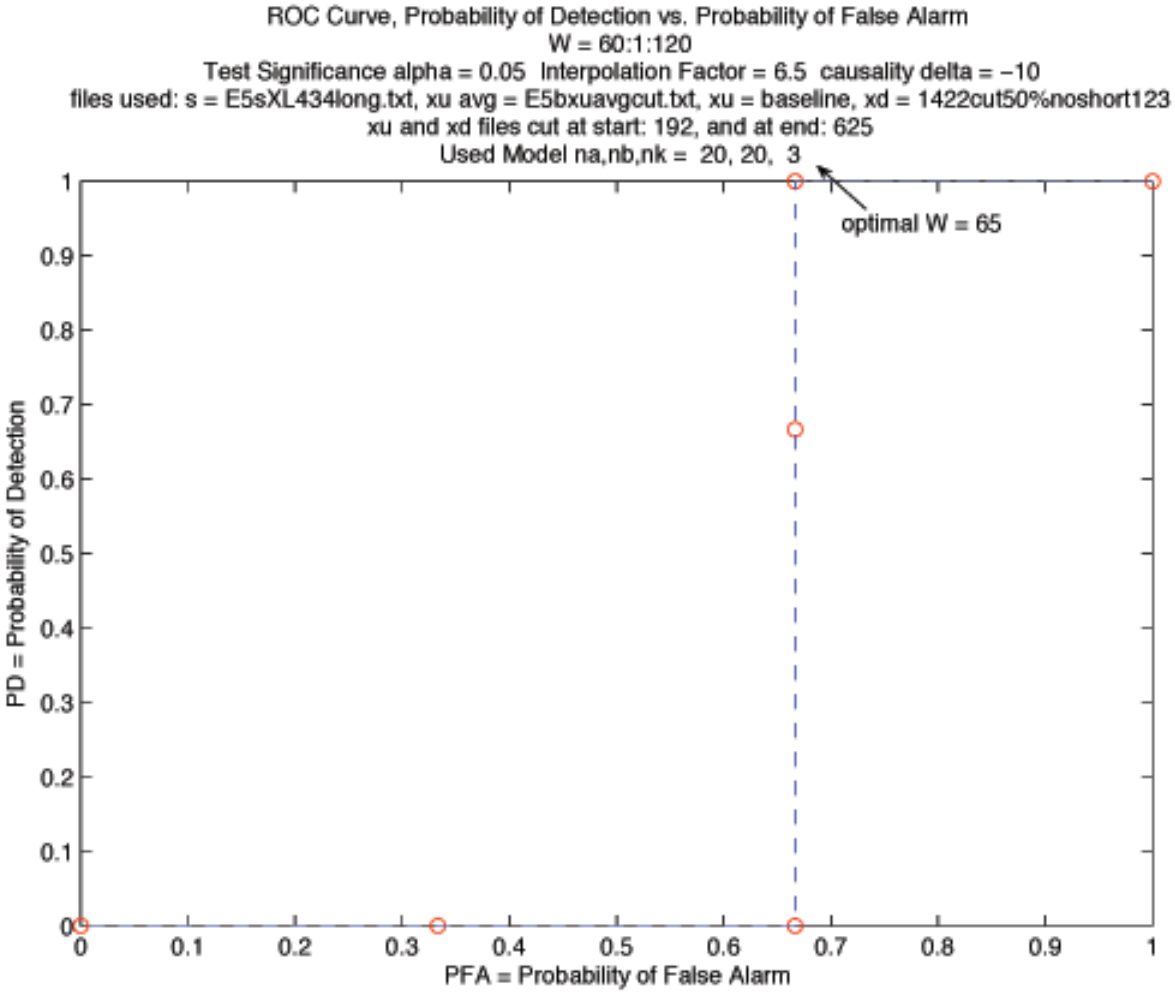


Figure 11.10: **ROC Curve for Case (3) with No TDR Birth Certificate Results: E5b ROC Curve, 1422-cut-90percent-noshort,7,8,9, No TDR Birth Certificate Case:** The detection performance is very poor. One could flip a coin and do as well. Future plans include reworking this experiment, explaining the suboptimal performance and making improvements as necessary.

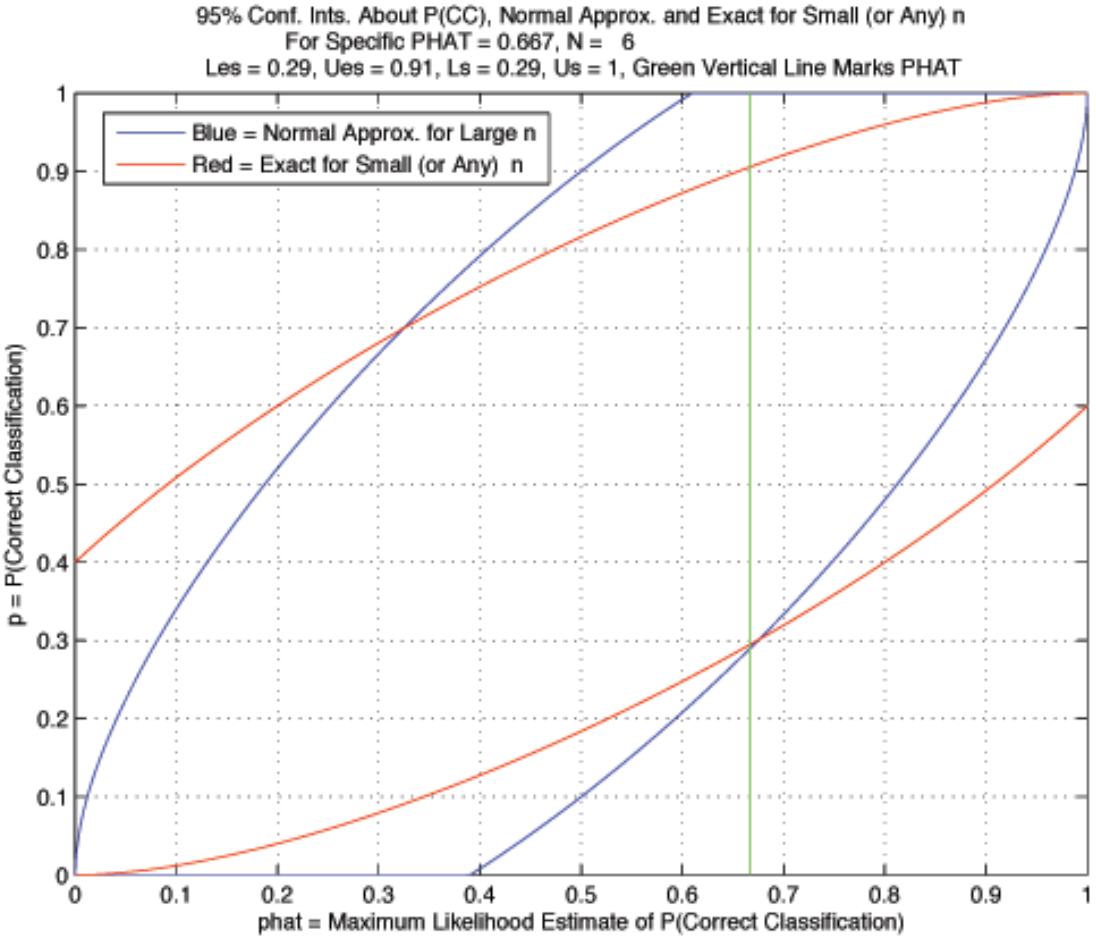


Figure 11.11: **Confidence Interval for Case (3) with No TDR Birth Certificate Results: E5b CI Both, 1403-cut-30percent-short, 3,4,5, No Birth Certificate Case:** The confidence interval plots are shown for both the Gaussian and exact cases. The green vertical line marks the calculated estimate of the probability of correct classification for the ROC operating point corresponding to $W = 65$ and $\widehat{p} = .667$.

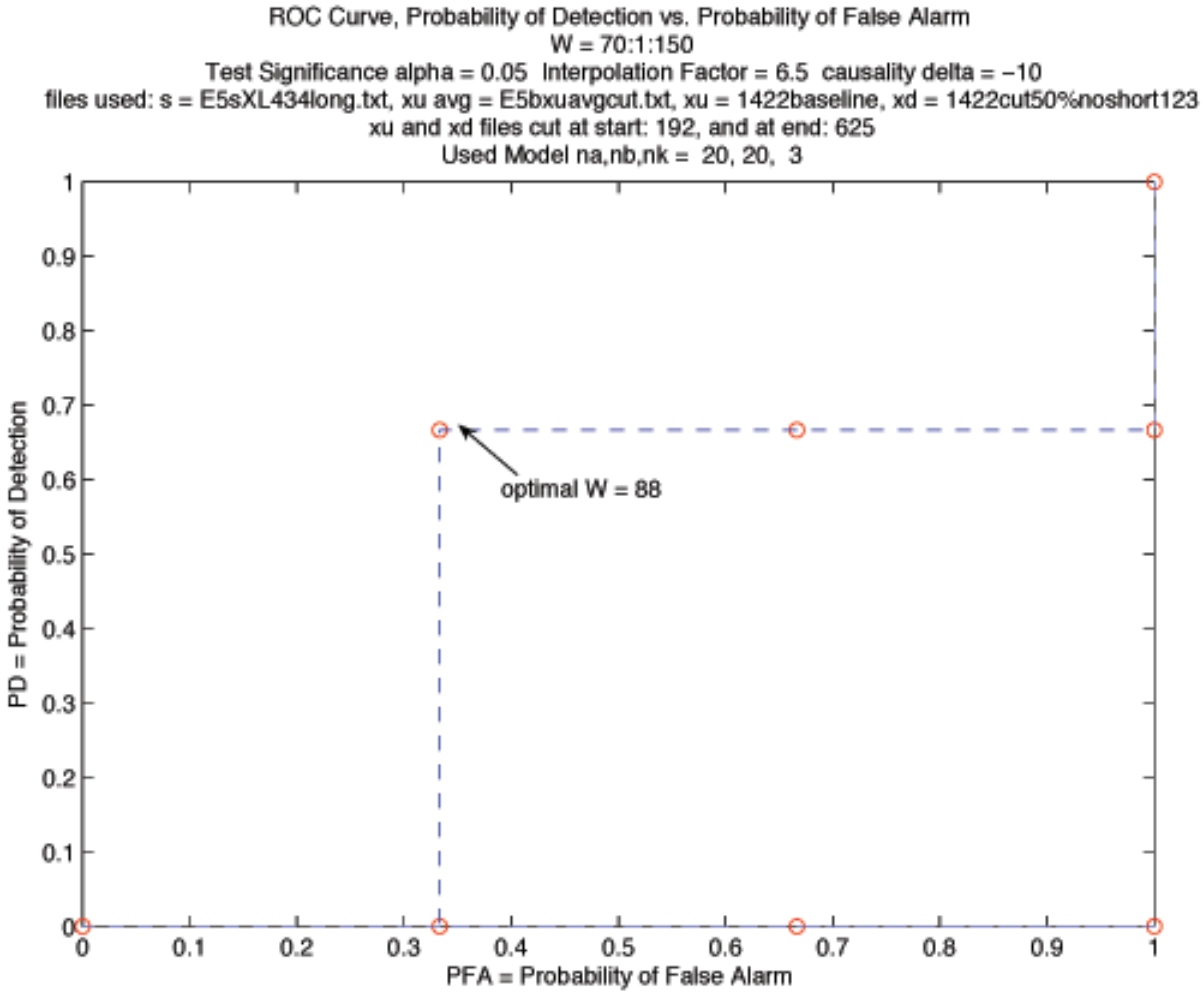


Figure 11.12: **ROC Curve for Case (3) with TDR Birth Certificate Results: E5b ROC Curve, 1422-cut-90percent-noshort,7,8,9, Birth Certificate Case:** The detection performance is very poor. Future plans include reworking this experiment, explaining the suboptimal performance and making improvements as necessary.

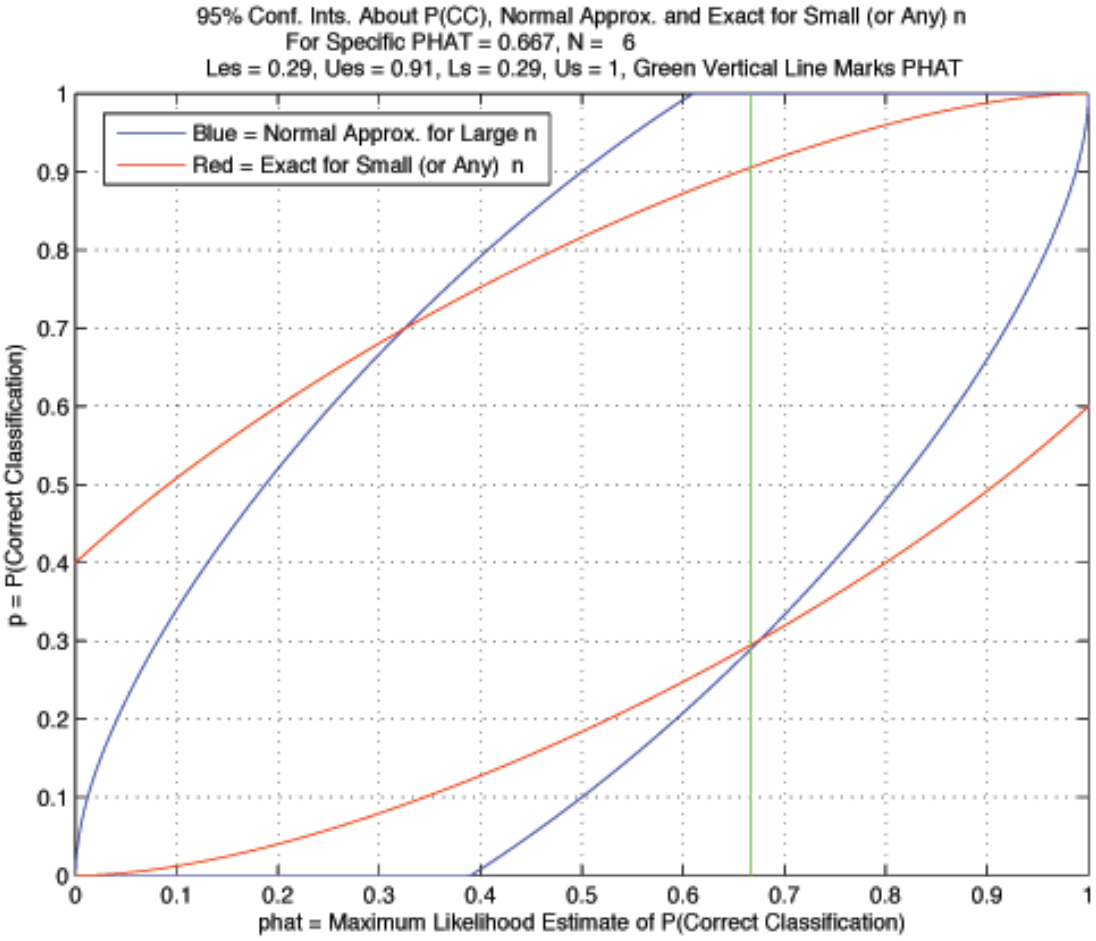


Figure 11.13: **Confidence Interval for Case (3) with TDR Birth Certificate Results: E5b CI Both, 1403-cut-30percent-short, 3,4,5, Birth Certificate Case:** The confidence interval plots are shown for both the Gaussian and exact cases. The green vertical line marks the calculated estimate of the probability of correct classication for the ROC operating point corresponding to $W = 88$ and $\widehat{p} = .667$.

Chapter 12

Conclusions

The model-based damage detection algorithms have been shown to perform as expected (well) for some representative examples of real TDR signals acquired using the 2D mockup fixture. We tested various damage types, including cuts, crushes and pinholes. See Experiments E1 and E5 in Chapters 10-12. The most difficult challenge in the project has been establishing sufficiently repeatable measurements in the test fixtures. The work required to achieve this repeatability was much more than expected at the beginning of the project and consumed a large portion of our time and resources. As a result, the number of cable damage experiments and analyses was smaller than planned. Further analysis is planned for future work.

The detection algorithms include a variety of tuning parameters that can be adjusted to affect performance in significant ways. This projected ended before it was possible to do serious tuning studies, but it is believed that there exists a great deal of latitude for tuning. Future work includes tuning studies.

12.1 Performance for Various Damage Types

Among the damage types, we tested various scenarios, as judged by visual inspection of the TDR signals. These include obvious damage, moderately obvious damage and subtle damage.

12.1.1 Case (1) Damage Resulting in a Short Circuit: 1403-cut-30percent-short, 3,4,5: TDR Birth Certificate, No TDR Birth Certificate Cases

If the damage causes a short circuit between the cable conductors, then the damage is generally obvious from visual inspection of the TDR signal, and damage detection algorithm performance is generally excellent (ideal). Examples include the cases demonstrated in this report for severe cuts and pinholes. See (Experiment E1, Major Damage), (Experiment E5, Case (1) Damage Resulting in a Short Circuit).

12.1.2 Case (2) Damage Not Resulting in a Short Circuit:1422-cut-90percent-noshort,7,8,9: TDR Birth Certificate, No TDR Birth Certificate Cases

If the damage does not cause a short circuit between the cable conductors, and the damage is moderately obvious by visual inspection of the TDR signal, then the damage detection algorithm performance is generally fair to poor. Examples include the cases demonstrated in this report for crushes and partial cuts. See (Experiment E5, Case (2) Moderate Damage), (Experiment E1, Minor 1 Damage) and (Experiment E1, All Damage). We expect that future

work would improve this performance. The algorithms contain several "tuning "parameters that are chosen by the user. We believe that with further time, data and resources, the algorithms could be tuned to provide much better performance.

12.1.3 Case (3) Damage Not Resulting in a Short Circuit: 1422-cut-50percent-noshort,1,2,3: TDR Birth Certificate, No TDR Birth Certificate Cases

If the damage does not cause a short circuit and the damage appears subtle by visual inspection of the TDR signal, then the damage detection performance is generally poor. Examples include the cases demonstrated in this report for crushes and partial cuts. See (Experiment E5, Case (3) Subtle Damage) and (Experiment E1, Minor 2 Damage).

In the case of (Experiment E1, Minor 3 Damage), the damage was very difficult to discern by visual inspection of the TDR signal, but the algorithms detected it with ideal performance. This is promising.

We expect that future work would improve this performance. The algorithms contain several "tuning "parameters that are chosen by the user. We believe that with further time, data and resources, the algorithms could be tuned to provide better performance.

12.2 Repeatability Issues

This project has led to the following specific conclusions regarding repeatability:

(a) Once the cable is attached (inserted into the fixture and connected to the pulse insertion unit (PIU)), the TDR signal has very little variability from measurement to measurement.

(b) In the 2D fixture, from attachment to attachment for a single cable, the variability is significant, but reasonably small.

(c) From cable to cable, the variability is even greater, and presents a significant issue. In fact, the variability from *cable to cable* is the greatest barrier to detection sensitivity (the "long pole in the tent") for the damage detection algorithms. The detection algorithms are very sensitive to small differences between the reference standard signal(s) and the signal(s) under test. However, when the measurements are not repeatable, the detection algorithms must be tuned to be less sensitive, so they do not mistake the variability in the measurements for damage. This is demonstrated in the receiver operating characteristic (ROC) curves depicted earlier.

In an ideal application, each TDR signals for each cable would be measured at the time the cable was installed in the device. This would provide a "TDR birth certificate" for each cable, against which one could compare the TDR signals measured at a later time. This would obviate the issues of cable-to-cable variability. For our application, without such "TDR birth certificates," we must live with the cable-to-cable variability and accept the resulting limits to probability of correct classification.

Chapter 13

Proposed Future Work

As with any project, it is possible to improve the damage detection system and/or apply it to other problems. Some of these are summarized in this section.

13.1 Additional Tests with More Damage Scenarios

We were able to test a set of damage types and damage scenarios during this study. However, future work would include expanding the number and types of damage scenarios. We would also attempt to obtain additional cables and test them to create a larger statistical sample size.

13.2 Work to Improve Repeatability of Measurements

13.2.1 The 2D Fixture

We spent an enormous fraction of our resources working to ensure that the measurements are reasonably repeatable. The main issue is making the 2D fixture have characteristics that do not vary from test to test. One issue was the PIU holding fixture. Improving the PIU holding fixture could further enhance the repeatability of the measurements. The Teflon arms that hold the PIU are flexible, and allow movement that may affect the repeatability of measurements. These arms could be replaced by more rigid metal arms. In addition, having the two arms redesigned to to form a single PIU holding fixture would prevent relative movement. The new holding fixture should also ensure that the PIU placement relative to the DUT is exact and the measurements are more repeatable.

13.2.2 The Pulse Insertion Unit (PIU)

The pulse insertion unit turned out to be the system component on which we spent the most time and effort. The PIU determines how well the cable impedance is matched, so the TDR signals are very dependent on the PIU and its design. A worthwhile task would be to contract with Picosecond Labs to redesign the PIU for improved repeatability.

13.3 Tuning Studies

The detection algorithms include a variety of tuning parameters that can be adjusted to affect performance in significant ways. This project ended before it was possible to do comprehensive tuning studies, but it is believed that there exists a great deal of latitude for tuning. The tuning parameters include the following: (1) *The sample sizes of the damage classes:* Larger sample sizes would add great value to the analysis. (2) *The WSSR window length W:* The ROC curves are parameterized by W. (3) *The statistical significance α of the confidence interval tests used for whiteness testing and WSSR testing:* In this study, only $\alpha = .95$ was used. We have the ability to choose a wide variety of values for α. (4) *Model parameters (N_a, N_b, N_k).* (5) *Priors and Costs (see the next section):* These can have a dramatic effect upon the ROC curves. (6) *Decimation factor M for the preprocessing steps:* This can affect the model order and the ability to obtain white innovations. (7) *Interpolation factor L for detection:* This factor can have significant effect on the resolution with which we can tune the WSSR window width W. Also, the derivation of the WSSR tests assumed a large sample size (large number of samples in the innovations signal), and interpolation factor has a direct effect on our ability to satisfy that assumption.

13.4 Study to Include Prior Probabilities and Costs for the Various Cable Damage Conditions

The Bayesian detection algorithm performance measures include the possibility of folding in any prior knowledge about the prevalence of various cable damage conditions. This can be done by estimating costs and prior probabilities, or "priors"of the various cable conditions based upon experimental data and results. These costs and priors can then be used in choosing decision thresholds for the ROC curves and in the formula for probability of correct classification. The choices of these parameters can have very significant effects on the performance calculations. For the work reported here, the priors were not known, so we assumed that they were equal. This was done by setting $P(H_0) = P(H_1) = 1/2$. This means that we assume that the null hypothesis H_0 for which the cable is not damaged is equally as likely as hypothesis H_1 for which the cable is damaged. In the notation of the chapter on detection performance measuremnet, the decision costs were assumed to be $C_{00} = C_{11} = 0$ and $C_{01} = C_{10} = 1$.

For the cable damage detection problem, we have reason to believe that these prior assignments are not representative of the population of cables we have observed. Further study with a large number of cables in various damage scenarios would allow us to create better estimates of the priors and thus better estimates of detection algorithm performance. We can also estimate decision costs that are more representative of the actual problem.

13.5 Alternative Detection Algorithms and Data Fusion

It is possible to create and evaluate the performance of multiple detection algorithms. For example, one could adapt the confidence interval test used for repeatability assessment to make it a damage detector. Another scheme is to use a variation on the signal average residuals defined as $r(n) = x_D(n) - x_{Uavg}(n)$, where $x_{Uavg}(n)$ is a sample ensemble average of a set of known reference cable signals (see the chapter on model-base signal processing). We actually implemented this scheme and ran one simple test case. We found that the WSSR is able to detect obvious cable damage when applied to the signal average residual. However, more work would need to be done to assess performance under more challenging scenarios.

Once ROC curves are created for all of the various detection schemes, we can use Bayesian decision-level fusion algorithms to create a fused decision based upon all of the detection modalities. A composite ROC curve can then be created for the overall fused system [28, 29, 30, 31, 32, 33, 34, 35].

13.6 Advanced Bootstrap Algorithms for Repeatability Tests

We implemented some basic bootstrap algorithms for estimating the confidence interval on the mean TDR signal when the sample size is small. Future work would include implementing and testing some more advanced bootstrap algorithms that promise even better accuracy.

13.7 Extensions to the ROC and Confidence Interval Algorithms

The ROC curve algorithms can be extended to include the confidence interval bounds directly on the ROC curve. G. Clark has derived the equations to effect this result, but has not implemented them in the software.

The confidence intervals used for the zero-mean whiteness test, the WSSR and the confidence interval on the probability of correct classification could be extended to the small sample size case using bootstrap algorithms.

13.8 Work with Other Cables/Applications

The ECATS system and the algorithms could be applied to other applications at LLNL and elsewhere. These include any new device systems that might be developed in the future, including the proposed RRW.

The system could also be applied to optical cable damage detection at LLNL NIF or a SNLL. G. Clark has done preliminary groundwork by discussing possible work with both NIF and SNLL.

13.9 Safety Improvements

Safety was primary driver for the initial test design [7]. The TDR tester is very safe and was approved by the LLNL High Explosives Instrument Committee (HEIC) for use with certain explosive components (see HEIC memo WESG-2006-06-01). However, some safety improvements may be required prior to use in nuclear explosive operations. Specifically, better separation of operational and safety functional components will simplify and enhance the safety theme. In addition, improved mechanical barriers to protect safety circuitry are required. These changes will result in better compliance with the safety guidance for electrical testers used in nuclear explosive operations (DG10001, EP401075 and DOE order 452.2C) and facilitate Nuclear Explosive Safety Group (NESSG) approval.

Bibliography

[1] G. A. Clark, *Cable Damage Detection Algorithms for Time Domain Reflectometry Signals*, Engineering Poster Session, Lawrence Livermore National Laboratory, February 18, 2009.

[2] G. A. Clark, *Cable Damage Detection Algorithms for Time Domain Reflectometry Signals*, Signal and Imaging Conference, Center for Advanced Signal and Imaging Sciences, Lawrence Livermore National Laboratory, LLNL report UCRL-CONF-408841, UCRL-ABS-407974, November 20-21, 2008.

[3] G. A. Clark, *Model-Based Algorithms for Detecting Damage in Ultrasonic Nondestructive Evaluation Measurements*, Invited Lecture, Special Session on Model-Based Signal Processing (Joint Signal Processing & Acoustical Oceanography), Invited by Prof. Christian Pichot (U. Nice, France), 155th Meeting of the Acoustical Society of America, 5th Forum Acusticum of the EAA, 9th Conres Francais DAcoustique of the SFA, Integrating 7th EUROpean Conference on NOISE Control Euronoise, 9th European Conference on Underwater Acoustics, ECUA, 60th Anniversary of the SFA, Paris, France, Lawrence Livermore National Laboratory report LLNL-CONF-405008, June 29-July 4, 2008.

[4] G. A. Clark, *Cable Damage Detection Using Time Domain Reflectometry and Model-Based Algorithms*, Sixth Annual Sensors Workshop 2008, Lawrence Livermore National Laboratory (LLNL), LLNL-ABS-401926 and LLNL-CONF-402567, April 1-2, 2008.

[5] G. A. Clark, *Model-Based Algorithms for Detecting Cable Damage from Time-Domain Reflectometry Measurements*, Signal and Imaging Conference, Center for Advanced Signal and Imaging Sciences, Lawrence Livermore National Laboratory, LLNL report UCRL-CONF-236598, UCRL-ABS-235724, November 15-16, 2007.

[6] G. A. Clark, *Model-Based Algorithms for Detecting Cable Damage from Time-Domain Reflectometry Measurements*, Seventh Biennial Tri-Laboratory Engineering Conference, Albuquerque, N.M., Lawrence Livermore National Laboratory report IM No. 347234, UCRL-CONF-230752, UCRL-ABS-229523, May 7-10, 2007.

[7] G. A. Clark and E. F. Breitfeller, *Model-Based Algorithms for Detecting Cable Damage from Time Domain Reflectometry Measurements*, Signal and Imaging Sciences Workshop 2005, Lawrence Livermore National Laboratory (LLNL), LLNL-ABS-215990 and LLNL-CONF-217125, November 17-18, 2005.

[8] Robbins, C. L., *Operating Procedure for the TDR Tester - PT4194 (U)*, Lawrence Livermore National Laboratory Document AF-PT4194 rev D, released via FCO20081777LL, October 13, 2008.

[9] Robbins, C.L. and Morey, R. M., *Nuclear Explosive Tester Safety Specification, PT-4194 Time Domain Reflectometry Tester (U)*, Lawrence Livermore National Laboratory Document TS-PT4194 rev A, released via AER20081776LL, December 4, 2008.

[10] C. L. Robbins and Morey, R. M., *ECATS: Electromagnetic Cable Assembly Tester System*, Lawrence Livermore National Laboratory Report UCRL-VIDEO-229185, PR: 005086, March 14, 2007.

[11] R. Morey, C. L. Robbins, G. A. Clark, C. Smith, A. Agoston, *ECATS: Electromagnetic Cable Assembly Tester System, A Safe, High Fidelity Tool for Transmission Lines*, Lawrence Livermore National Laboratory Report UCRL-MI-229180, R&D100 Entry, 2007.

[12] C. L. Robbins, G. H. Thomas, R. Morey, G. A. Clark, P. Souza, T. Le, J. Parker, J. A. Jackson, K. Lu, *Time Domain Reflectometry for Evaluating Detonator Cable Assemblies*, Lawrence Livermore National Laboratory report UCRL-POST-225062, National Nuclear Security Administration Future Technologies Conference II, Washington, D.C., October 11-13, 2006.

[13] G. A. Clark, M. C. Axelrod and D. D. Scott, *Classification of Heart Valve Sounds from Experiments in an Anechoic Water Tank*, Lawrence Livermore National Laboratory report, UCRL-JC-134634, June 1, 1999.

[14] G. A. Clark, *Impulse Response Estimation for Spatial Resolution Enhancement in 3D Ultrasonic Nondestructive Evaluation Imagery,* Proceedings of the Signal and Imaging Sciences Workshop, Center for Advanced Signal and Imaging Sciences, Lawrence Livermore National Laboratory,UCRL-ABS-207826, UCRL-PRES-208121, November 19, 2004.

[15] G. A. Clark, *Impulse Response Estimation for Spatial Resolution Enhancement in Ultrasonic Nondestructive Evaluation Imaging,* Lawrence Livermore National Laboratory Report UCRL-TR-205101, June 25, 2004.

[16] G. A. Clark, D.M. Tilly, and W. D. Cook, *Ultrasonic Signal/Image Restoration for Quantitative NDE,* NDT International, Vol. 19, No. 3, June 1986.

[17] G. A. Clark and E. J. Bogdan, *Summary of Rules for Sampling, Decimation and Interpolation of Experimental Data,* Lawrence Livermore National Laboratory Report UCRL-19411, April 1982.

[18] J. V. Candy, G. A. Clark, and D. M. Goodman, *Transient Electromagnetic Signal Processing: An Overview of Techniques,* Chapter in the book, *Time Domain Measurements in Electromagnetics,* edited by Edmund K. Miller, Van Nostrand Reinhold Co., New York, 1986, pp. 416-455.

[19] G. A. Clark and F. L. Barnes, "Ultrasonic Signal Resolution Enhancement for Adhesive Thickness Measurements," Review of Progress in Quantitative NDE, Williamsburg, VA., June 21-28, 1987.

[20] G. A. Clark, S. E. Benson, and D. K. Lewis, *Characterization of Wideband Ultrasonic Transducers Using Pulse Preshaping,* Review of Progress in Quantitative NDE, Williamsburg, VA., June 21-28, 1987.

[21] A. S. Willsky, *A Survey of Design Methods for Failure Detection in Dynamic Systems*, Automatica, Vol. 12, 1976.

[22] T. W. Anderson *An Introduction to Multivariate Statistical Analysis*, Wiley, New York, 1958.

[23] F. Schweppe, *Uncertain Dynamic Systems*, Prentice-Hall, Englewood Cliffs, N.J., 1973.

[24] R. K. Mehra and J. Peschon, *An Innovations Approach to Fault Detection and Diagnosis in Dynamic Systems*, Automatica, Vol. 7, 1971.

[25] A. M. Zoubir and D. R. Iskander, *Bootstrap Techniques for Signal Processing*, Cambridge U. Press, ISBN-13 978-0-521-03405-0, 2004.

[26] A. M. Zoubir and B. Boashash, *The Bootstrap and its Application in Signal Processing*, IEEE Signal Processing Magazine, January 1998.

[27] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, Chapman & Hall/CRC, 1998.

[28] B. Efron and R. Tibshirani, *Cross-Validation and the Bootstrap: Estimating the Error Rate of a Prediction Rule*, Unknown citation, circa 1993.

[29] R. P. S. Mahler, *Statistical Multisource-Multitarget Information Fusion*, Artech House, Inc., ISBN 13:978-1-59693-092-6, 2007.

[30] L. A. Klein, *Sensor and Data Fusion*, SPIE - The International Society for Optical Engineering, ISBN 0-8194-5435-4, 2004.

[31] R. T. Antony, *Principles of Data Fusion Automation*, Artech House, ISBN 0-89006-760-0, 1995.

[32] B. V. Dasarathy, *Decision Fusion*, IEEE Computer Society Press, ISBN 0-8186-4452-4, 1994.

[33] E. Waltz, J. Llinas, *Multisensor Data Fusion*, Artech House, ISBN 0-89006-277-3, 1990.

[34] J. J. Clark, A. L. Yuille, *Data Fusion for Sensory Information Processing Systems*, Kluwer, ISBN 0-7923-9120-9, 1990.

[35] X. E. Gros, *NDT Data Fusion*, Arnold, Wiley, 1997.

[36] X. E. Gros, *Applications of NDT Data Fusion*, Kluwer, 2001.

[37] G. A. Clark, *Practical Aspects of Sampling and the Discrete Fourier Transform,* LLNL Continuing Education Series, November 1983, with ongoing presentations over the past two decades.

[38] G. A. Clark, *Elements of Digital Signal and Image Processing - Practical and Theoretical Preparation for MATLAB Training,* Electronics Engineering Dept., Defense Sciences Engineering Division, Lawrence Livermore National Laboratory, UCRL-PRES-154699, September 2003.

[39] S. Treitel and E. A. Robinson, *The Design of High-Resolution Digital Filters,* IEEE Trans. Geoscience Electronics, Vol. GE-4, No. 1, 1966.

[40] E. A. Robinson and S. Treitel *Geophysical Signal Analysis,* Prentice Hall, 1980.

[41] R. M. Gray, *On the Asymptotic Eigenvalue Distribution of Toeplitz Matrices*, IEEE Trans. Information Theory, IT-18, pp. 725-730, 1972.

[42] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, Prentice Hall, 1998.

[43] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume II: Detection Theory*, Prentice Hall, 1998.

[44] J. V. Candy, *Model-Based Signal Processing*, IEEE Press and Wiley-Interscience, ISBN-13 978-0-471-23632-0 (cloth), ISBN-10 0-471-23632-2 (cloth), 2006.

[45] J. V. Candy, *Signal Processing: The Modern Approach,* McGraw-Hill, New York, NY, 1988.

[46] J. V. Candy, *Signal Processing: The Model-Based Approach,* McGraw-Hill, New York, NY, 1986.

[47] H. W. Sorenson, *Parameter Estimation,* Marcel Decker, Inc., 1980.

[48] T. Kailath, *Lectures on Linear Least-Squares Estimation*, Spinger-Verlag, 1976.

[49] A. Papoulis, *Probability, Random Variables and Stochastic Processes,* McGraw-Hill, 1965.

[50] A. Jazwinski, *Stochastic Processes and Filtering Theory,* Academic Press, New York, NY, 1970.

[51] H. L. Van Trees, *Detection,Estimation and Modulation Theory,* John Wiley and Sons, 1968.

[52] A. D. Whalen, *Detection of Signals in Noise,* Academic Press, New York, 1971.

[53] L. Ljung, *System Identification, Theory for the User*, Second Edition, Prentice Hall, 1999.

[54] L. Ljung, *System Identification, Theory for the User*, Prentice Hall, 1987.

[55] L. Ljung, *Theory and Practice of Recursive System Identification*, MIT Press, 1983.

[56] R. E. Crochiere and L. R. Rabiner, *Interpolation and Decimation of Digital Signals - A Tutorial Review*, Proc. IEEE, Vol. 69, No. 3, MArch 1981.

[57] Digital Signal Processing Committee of ASSP, *Programs for Digital Signal Processing*, IEEE Acoustics, Speech and Signal Processing Society, IEEE Press, 1979.

[58] A. V. Oppenheim and R. W. Schafer,*Digital Signal Processing*, Prentice-Hall, 1975.

[59] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, Prentice-Hall, 1975.

[60] A. Whalen, *Detection of Signals in Noise* Academic Press, 1971.

[61] C. D. McGillem and G. R. Cooper, *Continuous and Discrete Signal and System Analysis*, Holt, Rinehart and Winston, 1974.

[62] A. V. Oppenheim, A. S. Willsky and I. T. Young, *Signals and Systems*, Prentice-Hall, H.J., 1983.

[63] R. Bracewell, *The Fourier Transform and its Applications*, McGraw-Hill, 1965.

[64] R. E. Crochiere and L. R. Rabiner, *Multirate Digital Signal Processing*, Prentice-Hall, 1983.

[65] R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification*, Wiley, 2001.

[66] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley, 1973.

[67] R. V. Hogg and A. T. Craig, *Introduction to Mathematical Statisitics*, Macmillan Publishing Co., 1978.

[68] M. R. Spiegel, *Theory and Problems of Statistics, Schaum's Outline Series*, Schaum Pub. Co., New York, 1961.

[69] *MATLAB Reference Manual,* The Mathworks, Natick Massachussetts, 1993.

[70] *System Identification Toolbox*, The Mathworks, Natick Massachussetts, 2005.